

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using SW model

Run on: February 18, 2001, 19:19:51 ; Search time 3075.67 Seconds
(Without alignments)
6383.969 Million cell updates/sec

Title: US-09-143-828-3
Perfect score: 2802
Sequence: 1 tgaattcgtggtcgtcgtg.....aaaaaaaaaaaaaaaaaaaaa 2802

Scoring table: IDENTITY_MNC
Gapop 10.0 , Gapext 1.0

Searched: 7991742 seqs, 3503743858 residues
Total number of hits satisfying chosen parameters: 15983484

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*

1: gb.est1:*
2: gb.est2:*
3: gb.est3:*
4: gb.est4:*
5: gb.est5:*
6: gb.est6:*
7: gb.est7:*
8: gb.est8:*
9: gb.est9:*
10: gb.est10:*
11: gb.est11:*
12: gb.est12:*
13: gb.est13:*
14: gb.est14:*
15: gb.est15:*
16: gb.est16:*
17: gb.est17:*
18: gb.est18:*
19: gb.est19:*
20: gb.est20:*
21: gb.est21:*
22: gb.est22:*
23: gb.est23:*
24: gb.est24:*
25: gb.est25:*
26: gb.est26:*
27: gb.est27:*
28: gb.est28:*
29: gb.est29:*
30: gb.est30:*
31: gb.est31:*
32: gb.est32:*
33: gb.est33:*
34: gb.est34:*
35: gb.est35:*
36: gb.est36:*
37: gb.est37:*
38: gb.est38:*
39: gb.est39:*
40: gb.est40:*
41: em.estda:*
42: em.estfun:*
43: em.esthum1:*

44: em.esthum2:*
45: em.esthum3:*
46: em.esthum4:*
47: em.esthum5:*
48: em.esthum6:*
49: em.esthum7:*
50: em.esthum8:*
51: em.esthum9:*
52: em.esthum10:*
53: em.esthum11:*
54: em.esthum12:*
55: em.esthum13:*
56: em.esthum14:*
57: em.esthum15:*
58: em.esthum16:*
59: em.esthum17:*
60: em.esthum18:*
61: em.esthum19:*
62: em.esthum20:*
63: em.estin1:*
64: em.estin2:*
65: em.estin3:*
66: em.estin4:*
67: em.estrov1:*
68: em.estrov2:*
69: em.estp11:*
70: em.estp12:*
71: em.estp13:*
72: em.estp14:*
73: em.estp15:*
74: em.estrov1:*
75: em.estrov2:*
76: em.estrov3:*
77: em.estrov4:*
78: em.estrov5:*
79: em.estrov6:*
80: em.estrov7:*
81: em.estrov8:*
82: em.estrov9:*
83: em.estrov10:*
84: em.estrov11:*
85: em.estrov12:*
86: em.estrov13:*
87: gb.est41:*
88: gb.est42:*
89: gb.est43:*
90: gb.est44:*
91: gb.est45:*
92: gb.est46:*
93: gb.est47:*
94: gb.est48:*
95: gb.est49:*
96: gb.est50:*
97: gb.est51:*
98: gb.est52:*
99: gb.est53:*
100: gb.est54:*
101: gb.est55:*
102: gb.est56:*
103: gb.est57:*
104: gb.est58:*
105: gb.est59:*
106: gb.est60:*
107: gb.est61:*
108: gb.est62:*
109: gb.est63:*
110: gb.est64:*
111: gb.est65:*
112: em.esthum21:*
113: em.esthum22:*
114: em.esthum23:*
115: em.estom1:*
116: em.estom2:*

117: em_estp16:**
 118: em_estp17:**
 119: em_estp18:**
 120: em_estp19:**
 121: em_estp20:**
 122: em_estp21:**
 123: em_estp22:**
 124: em_estp23:**
 125: em_estp24:**
 126: em_estp25:**
 127: em_estp26:**
 128: em_estp27:**
 129: em_estp28:**
 130: em_estp29:**
 131: em_estp30:**
 132: em_estp31:**
 133: em_estp32:**
 134: em_estp33:**
 135: em_estp34:**
 136: em_estp35:**
 137: em_estp36:**
 138: em_estp37:**
 139: em_estp38:**
 140: em_estp39:**
 141: em_estp40:**
 142: em_estp41:**
 143: em_estp42:**
 144: em_estp43:**
 145: em_estp44:**
 146: em_estp45:**
 147: em_estp46:**
 148: em_estp47:**
 149: em_estp48:**
 150: em_estp49:**
 151: em_estp50:**
 152: em_estp51:**
 153: em_estp52:**
 154: em_estp53:**
 155: em_estp54:**
 156: em_estp55:**
 157: em_estp56:**
 158: em_estp57:**
 159: em_estp58:**
 160: em_estp59:**
 161: em_estp60:**
 162: em_estp61:**
 163: em_estp62:**
 164: em_estp63:**
 165: em_estp64:**
 166: em_estp65:**
 167: em_estp66:**
 168: em_estp67:**
 169: em_estp68:**
 170: em_estp69:**
 171: em_estp70:**
 172: em_estp71:**
 173: em_estp72:**
 174: em_estp73:**
 175: em_estp74:**
 176: em_estp75:**
 177: em_estp76:**
 178: em_estp77:**
 179: em_estp78:**
 180: em_estp79:**
 181: em_estp80:**
 182: em_estp81:**
 183: em_estp82:**
 184: em_estp83:**
 185: em_estp84:**
 186: em_estp85:**
 187: em_estp86:**
 188: em_estp87:**
 189: em_estp88:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	523.2	18.7	638	AV651714	AV651714 AV651714
2	445.2	15.9	461	AA699679	AA699679 z178c07.s
3	416	14.8	454	N94132	N94132 za25906.r1
4	340	12.1	343	AI248626	AI248626 qb77b12.x
5	338.2	12.1	794	AI746915	AI746915 u109a10.y
6	335	12.0	365	N69311	N69311 za25906.s1
7	201	7.2	472	AA277370	AA277370 va81a12.r
8	199.6	7.1	493	AM871811	AM871811 da94c06.y
9	185.4	6.6	200	AA679591	AA679591 z149d12.s
10	166.4	5.9	463	AM511148	AM511148 h422a05.x
11	165.4	5.9	215	AV108557	AV108557 AV108557
12	156.8	5.6	432	AM158294	AM158294 za39e07.x
13	154.4	5.5	601	AM107536	AM107536 u191a06.y
14	143.8	5.1	506	AI768052	AI768052 w146b06.x
15	140.8	5.0	1053	AM012320	AM012320 um07d07.y
16	132.8	4.7	540	BE573696	BE573696 60133441
17	132	4.7	463	AI049239	AI049239 uc85f08.y
18	118	4.2	724	AI046425	AI046425 uc85f04.x
19	108.4	3.9	468	AA396982	AA396982 mx86a06.r
20	102	3.6	463	AA107961	AA107961 mp04d08.r
21	101.2	3.6	506	AA679391	AA679391 z129a02.s
22	100.8	3.6	616	AI238397	AI238397 GH14435.5
23	99.2	3.5	732	AI393915	AI393915 AJ393915
24	99	3.5	370	AI024672	AI024672 ov60h04.x
25	97.6	3.5	690	BE266478	BE266478 LD08b66.5
26	97.6	3.5	1105	BE10116	BE10116 601302405
27	97.6	3.5	589	BE275249	BE275249 601122072
28	94	3.4	851	BE877201	BE877201 601485121
29	90.2	3.2	708	BE390042	BE390042 601285233
30	88.6	3.2	692	BE791424	BE791424 601582442
31	88.6	3.2	392	AM505439	AM505439 UI-HF-BNO
32	88.6	3.2	475	BE664197	BE664197 148439.MA
33	86.6	3.1	539	AI098643	AI098643 uc32f10.y
34	86.6	3.0	623	BE303050	BE303050 ba72h10.y
35	83.6	3.0	788	BE15476	BE15476 601664976
36	83.4	3.0	665	BE390515	BE390515 601284134
37	82.8	2.9	467	AA570939	AA570939 v185f05.r
38	82.4	2.9	959	BE903117	BE903117 601677193
39	81.8	2.9	564	AV609965	AV609965 AV609965
40	81.4	2.9	771	BE385638	BE385638 601278159
41	81.2	2.9	248	AA272757	AA272757 vb09h12.r
42	81	2.9	514	AA124058	AA124058 mo29c08.r
43	80.4	2.9	667	AM231123	AM231123 uc39g12.y
44	80.4	2.9	563	BE233303	BE233303 139430.MA
45	80	2.9	105	BE233303	BE233303 139430.MA

ALIGNMENTS

RESULT 1
 AV651714
 LOCUS
 DEFINITION AV651714 G1C Homo sapiens cDNA clone G1CSCG03 3', mRNA sequence.
 AV651714
 VERSION
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens

Qy	2210	taagtaaaaaaacaaacagaaacaaacaacaaatcttgatcaaaagagaaatgtaaagt	2269
Db	401	TAGTTAAAAAACAAACAGAAACAAACAACTTTGGATCAAAAGAG--AAATGATAGTG	343
Qy	2270	acaaaagcaagcacaaagaaatctccctctgtgtaagtcgtacgtctgtatgacggcactg	2329
Db	342	ACAAAGACAGACAAAGAAATTTCCCTGTGGAGTCTGAGCTGTATGCGGGGCACTGGG	283
Qy	2330	tacccaagtgaaggtctcccgaaagcacttgatctgtatgggcaagggcacaaactgcagct	2389
Db	282	TACCCAGTGAAGGTTCCGAGGACATGATCTGTATGGAGCAAGGCGCAACATCGACACT	223
Qy	2390	gtgagtcgctgtgtgtatttgltgaagtaagtcctgttcgcaactgatgagggcctg	2449
Db	222	GTGAGTGGCTGTGTGTATTTGGGTGATAGTAGTCTGTTGGCACTGATAGGGGCGCTGGG	163
Qy	2450	ttgttcctggggcctgtaatcctgtatgctctgtgacaaggtcaagcctgaacaatcagt	2509
Db	152	TTTGTCTCTGGCGCTGGAAATGCTGGGATCTGTGTGACAAACGTACCGCTGCACATCAGT	103
Qy	2510	taaacacccggagaaagaacacatttaacatgcacctatatctctgtgaacaatctatc	2569
Db	102	TAAACACACCGGAAMAAACATTATAGTACACTTATATTTCTGTGTACACATCTATTC	43
Qy	2570	tcaagctaaagggtgatgaagctgcctgcctgtttatatgcc	2611
Db	42	TCAAGCTAAAGGTTATGAAGTCCCTGCCTTGTATTATGCC	1

[illegible]

Eukaryote; Metazoa: Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
1 (bases 1 to 454)
Hillier, L., Clark, N., Dubouque, T., Elliston, K., Hawkins, M., Holman

TITLE	The Wasnu-Merck Est Project
JOURNAL	Unpublished (1995)
COMMENT	Contact: Wilson RK

This clone is available royalty-free through INL; contact the
Image Consortium (info@image.inl.gov) for further information
Seq primer: mob.MEGA-ET
High quality sequence stop: 446.
Email: estewatson.junsl.edu
Tel: 314 286 1800
Fax: 314 286 1810
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

FEATURES

```

/organism="Homo sapiens"
/db.xref="GDB:3801473"
/db.xref="taxon:9606"
/clone_image="293626"
/clone_id="Soares fetal liver spleen INFLS"
/sex="male"
/dev_stage="20 week post conception fetus"
/lab_host="DH10B (ampicillin resistant)"
/note="Organ: Liver and Spleen; Vector: p7T3D (Pharmacia)
with a modified polylinker. Site.1: Pac I; Site.2: Eco RI;
1st strand cDNA was primed with a Pac I oligo(dT) primer."
15, AACTGGAGAGATTAAATTAAAGACCTTTTATTTTTTTTTTTT 3'}

```

double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Pac I and cloned into the Pac I and Eco RI sites of the modified pTR73 vector. Library went through one round of normalization. Library constructed by Bento Soares and M.Fátima Bonaldo."

Query Match	14.8%	Score 416;	DB 143;	Length 454;
Best Local Similarity	98.4%;	Pred. No. 7.4e-86;		
Matches 441; Conservative	0;	Mismatches 5;	Indels 2;	Gaps 2.

QY 1802 agatcccaataagttcaagttgtgaaagggacccaagcgaccaagatatagccatctgg 1861

Db 7 AGATCCCAAAAGTGTCAAGGTGTGGAAGGACCAACCCACCAAGATGGGCCAATCTGG 20

Db 67 GGTCTATGCCACATATACCCACGTTTGTTCGCTTCCGAGTCTTTTCATTGCTACCTCTAA 126

QY 1922 tagtctgtccccaattcccaactgttcccccctctcttccgagctgctttgtgctca 1981

D_b 127 TAGTCCTGTCGCCACTTCCCACTCGTTCCCCCCTCCCTCTCTCCAGACGTGCATTGTGGGCTCC 186

187 AGCCCTCTACCTCATCCGAGCGCATGATCTCTGAGGAGTCTCCTTAGAGAGATGAGAA 246

QY 2042 gccagagagctgcaccaatgtcagaagcttgcatgacctcatctccggccacatcatt 2103

Db 247 TCCAGAGGCGTCCACCAATGTGAGAGCTTGGCAGACCTCATTTCCGCGCACATCATTT 308

[illegible]

QY 2162 gggtatacagcatt-gactcagatatagatcctg-agctcacagagtttatagtttaaaa 2219

Db 367 GGGTATACAGCAATTGGACTCAGATATAGATCCTGAAGCTCACAGAGTTTATAGTTAAAA 426

BY	4240	4241	4242	4243	4244	4245	4246	4247	4248	4249	4250	4251	4252	4253	4254	4255	4256	4257	4258	4259	4260	4261	4262	4263	4264	4265	4266	4267	4268	4269	4270	4271	4272	4273	4274	4275	4276	4277	4278	4279	4280	4281	4282	4283	4284	4285	4286	4287	4288	4289	4290	4291	4292	4293	4294	4295	4296	4297	4298	4299	4300	4301	4302	4303	4304	4305	4306	4307	4308	4309	4310	4311	4312	4313	4314	4315	4316	4317	4318	4319	4320	4321	4322	4323	4324	4325	4326	4327	4328	4329	4330	4331	4332	4333	4334	4335	4336	4337	4338	4339	4340	4341	4342	4343	4344	4345	4346	4347	4348	4349	4350	4351	4352	4353	4354	4355	4356	4357	4358	4359	4360	4361	4362	4363	4364	4365	4366	4367	4368	4369	4370	4371	4372	4373	4374	4375	4376	4377	4378	4379	4380	4381	4382	4383	4384	4385	4386	4387	4388	4389	4390	4391	4392	4393	4394	4395	4396	4397	4398	4399	4400	4401	4402	4403	4404	4405	4406	4407	4408	4409	4410	4411	4412	4413	4414	4415	4416	4417	4418	4419	4420	4421	4422	4423	4424	4425	4426	4427	4428	4429	4430	4431	4432	4433	4434	4435	4436	4437	4438	4439	4440	4441	4442	4443	4444	4445	4446	4447	4448	4449	4450	4451	4452	4453	4454	4455	4456	4457	4458	4459	4460	4461	4462	4463	4464	4465	4466	4467	4468	4469	4470	4471	4472	4473	4474	4475	4476	4477	4478	4479	4480	4481	4482	4483	4484	4485	4486	4487	4488	4489	4490	4491	4492	4493	4494	4495	4496	4497	4498	4499	4500	4501	4502	4503	4504	4505	4506	4507	4508	4509	4510	4511	4512	4513	4514	4515	4516	4517	4518	4519	4520	4521	4522	4523	4524	4525	4526	4527	4528	4529	4530	4531	4532	4533	4534	4535	4536	4537	4538	4539	4540	4541	4542	4543	4544	4545	4546	4547	4548	4549	4550	4551	4552	4553	4554	4555	4556	4557	4558	4559	4560	4561	4562	4563	4564	4565	4566	4567	4568	4569	4570	4571	4572	4573	4574	4575	4576	4577	4578	4579	4580	4581	4582	4583	4584	4585	4586	4587	4588	4589	4590	4591	4592	4593	4594	4595	4596	4597	4598	4599	4600	4601	4602	4603	4604	4605	4606	4607	4608	4609	4610	4611	4612	4613	4614	4615	4616	4617	4618	4619	4620	4621	4622	4623	4624	4625	4626	4627	4628	4629	4630	4631	4632	4633	4634	4635	4636	4637	4638	4639	4640	4641	4642	4643	4644	4645	4646	4647</
----	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	--------

AI248626/c	343 bp	MRNA	EST	01-DEC-1998
LOCUS				
AI248626:				

clone IMAGE:1850687 3' similar to gb:J03258 VITAMIN D3 RECEPTOR

ACCESSION	AI248626
VERSION	AI248626.1
	GI:3844023

SOURCE human.
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP)

JOURNAL
Unpublished (1997)
Detached Robert Straubhaar
ph r

Tel: (301) 496-1550
Email: Robert_Strausberg@nh.gov

IMAGE Consortium (info@image.lnl.gov) for further information
Insert Length: 608 Std Error: 0.00

seq primer: -400F 110M GAGCC
High quality sequence stop: 338.

Query Match 7.28; Score 201; DB 4; Length 472;
 Best Local Similarity 77.28; Pred. No. 4; 7e-36;
 Matches 294; Conservative 0; Mismatches 82; Indels 5; Gaps 4;

QY 1210 ccaatccctctctctccacagccgagctgtcgcagcac-cgcgctgtgtgagcag 1268
 |||||
 DB 1 CCAATCTCCCTCTTCCCAATGCTCTGCTGTCAGCCGACGCTNNNGTAGACCAA 60
 QY 1269 ctgcagagagcaatccgcatctacacgaagctcacatctgaatcgagccagcct 1328
 |||||
 DB 61 CTGACAGAGAGGTGGTGGCTCACCCTGAAGGCTACATGAGTGTATCGGCATATCCT 120
 QY 1329 gctcagagctctgtctcctgaagatcagctgtcagcagagctcgcagcatt 1388
 |||||
 DB 121 GCTCAGAGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 178
 QY 1389 gctcag 1448
 |||||
 DB 179 GCCCAGCAACCCAGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 238
 QY 1449 atgcag 1508
 |||||
 DB 239 ATGCAAGAGTATTATGACAGACAGATGCTGCTGCTGCTGCTGCTGCTGCTGCT 298
 QY 1509 gagcag 1568
 |||||
 DB 299 GGGGAGAGTATGACCAAGCA-CATCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCT 356
 QY 1569 gccagagagcagagagagagagagagagagagagagagagagagagag 1589
 |||||
 DB 357 ACTGATAGCCCAACACGCCCT 377

RESULT 8

AM871811 493 bp mRNA EST 22-JUN-2000
 LOCUS da94c06.y1 Xenopus laevis tadpole stage 24 Xenopus laevis cDNA
 DEFINITION clone IMAGE: 5' similar to gb:J03258 VITAMIN D3 RECEPTOR (HUMAN);
 mRNA sequence.

ACCESSION AM871811 GI:8005864
 VERSION AM871811.1
 KEYWORDS EST.
 SOURCE African clawed frog.

ORGANISM Xenopus laevis
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipiloidea; Pipidae;
 Xenopodidae; Xenopus.

REFERENCE 1 (bases 1 to 493)
 Clifton,S., Johnson,S.L., Blumberg,B., Song,J., Hillier,L., Page,D.
 , Martin,J., Wylie,T., Underwood,K., Theising,B., Bowers,Y., Person
 ,B., Gibbons,M., Harvey,N., Ritter,E., Jackson,Y., McCann,R.,
 Waterston,R. and Wilson,R.

TITLE Washu Xenopus EST project, 1999
 JOURNAL Unpublished (1999)
 COMMENT Other ESTs: da94c06.x1

CONTACT: Sandy Clifton, Ph.D.
 Washu Xenopus EST project, 1999
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810

EMAIL: est@wustl.edu
 Library constructed by B. Korn PhD. and S. Henze DNA Sequencing by:
 Washington University Genome Sequencing Center
 Clone distribution: Xenopus clones from this library are available
 through the Resource Center at the University of Maryland, Berlin, Germany (web
 address: www.rzpd.de)
 Seq primer: -408P from GIDCO
 High quality sequence stop: 418.

FEATURES

1..493
 Location/Qualifiers
 /organism="Xenopus laevis"
 /db_xref="taxon:8355"

/clone="IMAGE:"
 /clone_id="Xenopus laevis tadpole stage 24"
 /sex="mixed"
 /tissue_type="whole tadpole"
 /dev_stage="stage 24"
 /lab_host="DH10B"
 /note="Vector: pCMVSPORT6; Site_1: NotI; Site_2: SalI.
 Method of cloning used: directed ligation; InsertCheck:
 95.8 blue and 4.2 white. Library constructed by Dr.
 Bernhard Korn and Sabine Henze. Note: This is a Xenopus
 Gene Collection (XGC) library."

BASE COUNT 129 a 142 c 131 g 91 t
 Query Match 7.1%; Score 199.6; DB 96; Length 493;
 Best Local Similarity 68.1%; Pred. No. 1e-35;
 Matches 293; Conservative 0; Mismatches 134; Indels 3; Gaps 1;

QY 267 gcagatgag 326
 |||||
 DB 10 GAAG 69
 QY 327 tctcattcaatgcatgcatgcatgcatgcatgcatgcatgcatgcatgcatgcat 386
 |||||
 DB 70 TATCACTTCAATGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 129
 QY 387 cgcag 446
 |||||
 DB 130 AGGAGCTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 186
 QY 447 cggcag 506
 |||||
 DB 187 CGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 246
 QY 507 atgatatctcag 566
 |||||
 DB 247 TTGATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 306
 QY 567 gaacgag 626
 |||||
 DB 307 TTACAGAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 366
 QY 627 atcag 686
 |||||
 DB 367 CTCACCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 426

RESULT 9

AA679591 200 bp mRNA EST 19-DEC-1997
 LOCUS zj49d12.s1 Soares fetal_liver_spleen_INFIS.S1 Homo sapiens cDNA
 DEFINITION clone IMAGE:453623 3' similar to TR:6410518 G410518 ORPHAN NUCLEAR
 RECEPTOR OF STEROID/THYROID SUPERFAMILY. ; mRNA sequence.

ACCESSION AA679591 GI:2660113
 VERSION AA679591.1
 KEYWORDS EST.
 SOURCE human.

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homalidae; Homo.

REFERENCE 1 (bases 1 to 200)
 Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
 Kitzman,D., Kucada,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin
 ,J., Moore,B., Schellenberg,K., Stepecoe,M., Tan,F., Theising,B.,
 White,Y., Wylie,T., Waterston,R. and Wilson,R.
 Washu-NCI human EST Project
 Unpublished (1997)
 CONTACT: Wilson RK

JOURNAL
COMMENT

Unpublished (1999)
Contact: Chie Owa
Genome Science Laboratory
RIKEN
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
Tel: 81-298-36-9145
Fax: 81-298-36-9098
Email: genome-reserch.riken.go.jp

Thermolabile and thermostable DNA polymerases by
trehalose and its application for the synthesis of full length cDNA
(Proc. Natl. Acad. Sci. U.S.A. 95(2):520-524 (1998))
Transcriptional sequencing: A method for DNA sequencing using RNA
polymerase (Proc. Natl. Acad. Sci. U.S.A. 95(7):3455-3460 (1998))
Please visit our web site (<http://genome.riken.go.jp>) for
further details.

FEATURES
source

Location/Qualifiers
1. 215
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone_lib="2510047D01"
/clone_lib="Mus musculus liver C57BL/6J 13-day embryo"
/sex="mixed"
/tissue_type="liver"
/dev_stage="13-day embryo"

BASE COUNT 48 a 60 c 69 g 38 t
ORIGIN

Query Match 5.9%; Score 165.4; DB 31; Length 215;
Best Local Similarity 85.6%; Pred. No. 7e-28;
Matches 184; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

QY 340 tcatgacatgtgaagatgcaagagcttttcagagagagccatgaacagcagccgcgc 399
DB 1 TCGAGACCGCTGAGATGAGAGGCTTTTTCAGCAGAGCCCTGACACACATGTCGGC 60
QY 400 ttaggtgtcccttcggaagagcgctcgagatcacccggaagcccgagcagctgccc 459
DB 61 TGAGCTGCCCCCTTTCGAGAGGAACTGCGAGATCACCCGAGAGACGAGCGCACTGCC 120
QY 460 aggcctgcgcgcctgcgaagtgcttgagagcgagcagtgagagagagatgctgccc 519
DB 121 AGGCTGCTCCCTTTGCGAGAGTCTGAGATGAGAGAGAGAGATGATTAATGTCGC 180
QY 520 acgagagcctgtgagagagcgagccttgatcaag 554
DB 181 ATCGCGCTGTGAGAGAGCGGCGCTTGTATCAAG 215

RESULT 12
AM158294 432 bp mRNA EST 05-NOV-1999
LOCUS 2a39e07.x1 Xenopus EST library Xenopus laevis cDNA clone 2a39e07
DEFINITION 5' mRNA sequence.
ACCESSION AM158294.1 GI:6270323
VERSION EST.
KEYWORDS African clawed frog.
SOURCE Xenopus laevis
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
Xenopodinae; Xenopus.
1 (bases 1 to 432)
REFERENCE
AUTHORS Schutz,K., de la Basile,M., Huang,E.N., Nasclmento,L., Preston,R.,
Shah,R., Swaby,I., Shekher,M., Spiegel,L., Vill.M.D. and McComble
W.R.

TITLE Expressed sequence tags from Xenopus
JOURNAL Unpublished (1999)
COMMENT Contact: W. Richard McComble
Lita Annenberg Hazen Genome Sequencing Center
Cold Spring Harbor Laboratory
PO Box 100, Cold Spring Harbor, NY 11724, USA

Tel: 516 367 8884
Fax: 516 367 8874
Email: mcomble@csb.riken.go.jp
Plate: 2a39 row: e column: 07
Seq primer: M13 universal forward primer
High quality sequence stop: 432.
Location/Qualifiers
1. 432

FEATURES
source

/organism="Xenopus laevis"
/db_xref="taxon:8355"
/clone_lib="2a39e07"
/clone_lib="Xenopus EST library"
/tissue_type="total brain tissue"
/cell_line="W22-RGA"
/dev_stage="tadpole"
/note="Vector: Lambda Zap I; Site: 1: XbaI; This library
was supplied by Holly Cline (Cold Spring Harbor Labs).
cDNA synthesis with oligo dt Xba I (Xba I cloning site).
RNA: stage 50-56 tadpoles, total brain tissue, gtc
extraction method."

BASE COUNT 112 a 90 c 142 g 88 t
ORIGIN

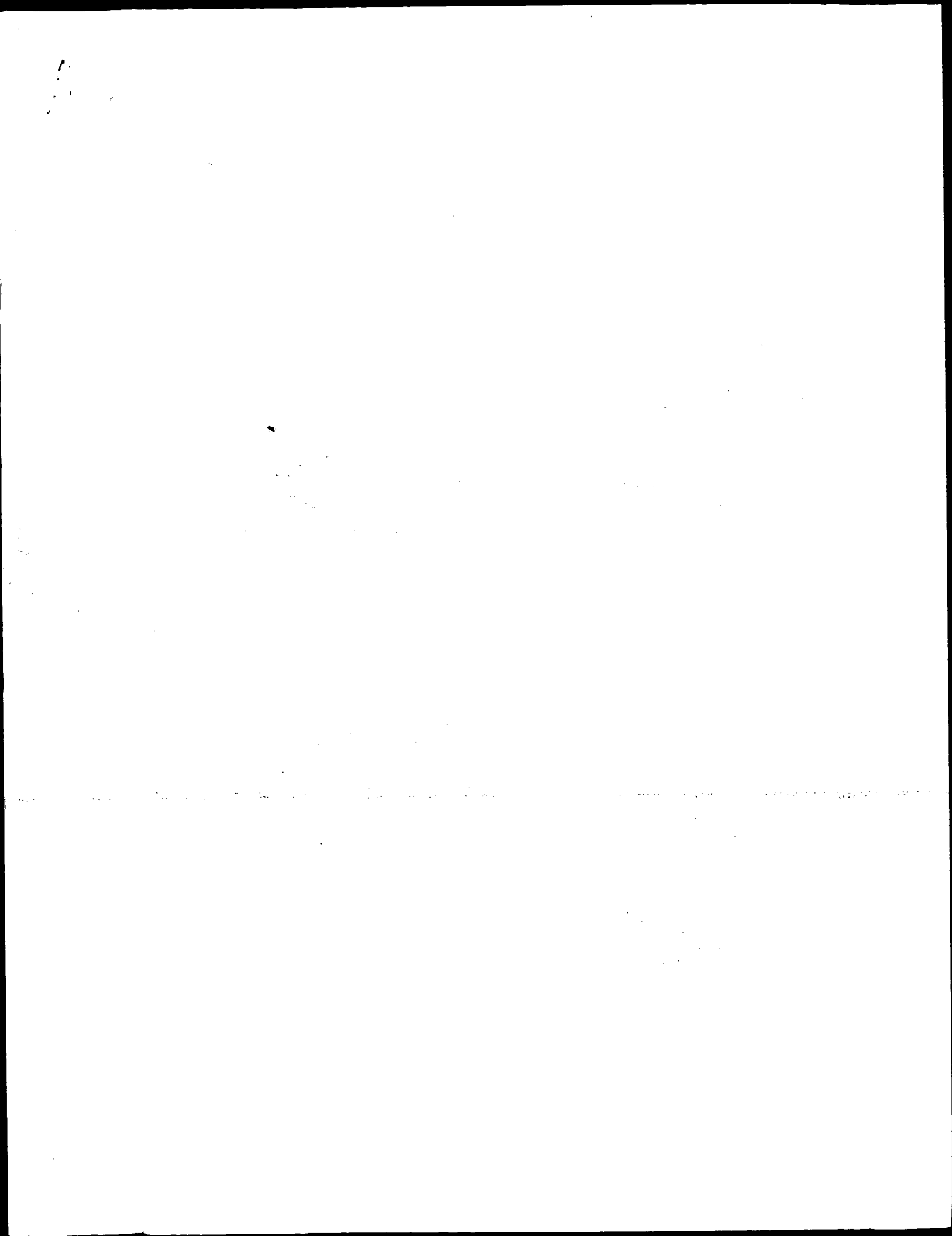
Query Match 5.6%; Score 156.8; DB 40; Length 432;
Best Local Similarity 74.0%; Pred. No. 7.9e-28;
Matches 213; Conservative 0; Mismatches 72; Indels 3; Gaps 1;

QY 267 gcagatgaggaagtcggaggtcccaaatctgcggtgtagtgaggacaaagccatgccc 326
DB 147 GAAAG 206
QY 327 tctacatcaatgtcatcatcatgtgagagatgagagagcttttcagagagagccatgaa 386
DB 207 TATCTCTCAATGCTATGACCTGCGAGGCTGCGAGAGAGATCTTACAGCGCGCTGAG 266
QY 387 cgcacagcgcgcgtgaggtgccccttcggaagagcgctgagatcacccggaagacc 446
DB 267 AGGAAGCTGCGGCTGAG 323
QY 447 cggcagagctgcagagctgcgccttcgcaagtgctgagagagcgagatgagagagag 506
DB 324 CGGCGCCACTGCTCGAGGCTGCTGCTCAAGAAATGCTGACATCGGCAAGAGAGAGAG 383
QY 507 atgatatgtctgcagagagcgctgtgagagagagagagagagagagagagagag 554
DB 384 TTGATCATGTCTGATGAG 431

RESULT 13
AM107536 601 bp mRNA EST 20-OCT-1999
LOCUS U191a06.y1 Suga mouse kidney mla Mus musculus cDNA clone
DEFINITION IMAGE:2159410 5' similar to gb:J03358 VITAMIN D3 RECEPTOR (HUMAN);
gb:U15548 Mus musculus beta 2 thyroid hormone receptor (MUSE);
mRNA sequence.
ACCESSION AM107536
VERSION AM107536.1 GI:6078336
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
AUTHORS Marra,M., Hillier,L., Kucaba,T., Martin,T., Beck,C., Wylie,T.,
Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person
B., Swaller,T., Gibbons,M., Page,D., Harvey,N., Schurk,R., Ritter
E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R.,
Waterston,R. and Wilson,R.

TITLE The WashU-NCI Mouse EST Project 1999
JOURNAL Unpublished (1999)
COMMENT Other-ESTs: U191a06.x1
Contact: Marra W/WashU-NCI Mouse EST Project 1999



GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 16, 2001, 16:49:16 ; Search time 124.29 seconds
(without alignments)
3533.206 Million cell updates/sec

Title: US-09-143-828-3

Perfect score: 2802
Sequence: 1 tgaattcggggccctgctg.....aaaaaaaaaaaaaaaaaaaaa 2802

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 280836 seqs, 80580151 residues

Total number of hits satisfying chosen parameters: 561672

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA.*
1: /cgn2_6/ptodata/2/1na/5A.CONB.seq.*
2: /cgn2_6/ptodata/2/1na/5B.CONB.seq.*
3: /cgn2_6/ptodata/2/1na/6.CONB.seq.*
4: /cgn2_6/ptodata/2/1na/PCTUS.CONB.seq.*
5: /cgn2_6/ptodata/2/1na/Backfile1.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	278.8	10.0	2043	1	US-07-737-736B-6
2	272.8	9.7	1399	1	US-07-737-736B-5
3	170.6	6.1	1450	1	US-08-459-489-1
4	170.6	6.1	1450	1	US-08-459-489-1
5	170.6	6.1	1450	1	US-07-843-350C-1
6	170.6	6.1	1450	4	PCT-US93-01559-1
7	97.6	3.5	1898	1	US-08-342-411A-1
8	97.6	3.5	1898	1	US-08-342-411A-1
9	97.6	3.5	1898	1	US-08-342-411A-1
10	97.6	3.5	2030	1	US-08-330-518-1
11	97.6	3.5	2030	2	US-08-330-283-1
12	97.6	3.5	2030	2	US-08-646-248-1
13	97.6	3.5	2030	4	PCT-US95-13924-1
14	97.6	3.5	1813	4	PCT-US94-12883-3
15	91.4	3.3	2928	2	US-08-095-728B-3
16	91.4	3.3	2928	2	US-08-095-728B-3
17	91.4	3.3	2940	5	5171671-1
18	91.4	3.3	3036	1	US-08-306-691B-52
19	91.4	3.3	3036	2	US-08-095-728B-1
20	91.4	3.3	3036	2	US-08-095-728B-1
21	91.4	3.3	3511	4	US-08-892-747-13
22	89.8	3.2	704	2	US-08-592-383-5
23	89.8	3.2	2940	2	US-08-592-383-5
24	85.2	3.0	1959	4	US-08-342-411A-3
25	85.2	3.0	1959	4	PCT-US94-12883-4
26	83.8	2.9	2658	2	US-08-592-383-3
27	80.8	2.9	1860	2	US-08-372-652-7
28	80.8	2.9	1860	4	PCT-US95-16311-7

29	79.6	2.8	1659	1	US-08-333-358-7	Sequence 7, Appl
30	79.6	2.8	1659	1	US-08-463-694-7	Sequence 7, Appl
31	79.6	2.8	1659	1	US-08-694-501-7	Sequence 1, Appl
32	77.6	2.8	2970	4	PCT-US92-06391-1	Sequence 2, Appl
33	77.6	2.7	1688	2	US-08-649-619B-2	Sequence 1, Appl
34	74.8	2.7	1649	2	US-08-466-130-1	Sequence 1, Appl
35	74.8	2.7	1649	2	PCT-US94-07266-1	Sequence 1, Appl
36	74.4	2.7	1576	5	5260432-1	Patent No. 5260432
37	74.4	2.7	2241	3	US-09-144-759-17	Sequence 17, Appl
38	74.4	2.7	2295	3	US-09-144-759-19	Sequence 19, Appl
39	74.4	2.7	2301	3	US-09-144-759-21	Sequence 21, Appl
40	73.2	2.6	2468	1	US-08-333-358-11	Sequence 11, Appl
41	73.2	2.6	2468	1	US-08-463-694-11	Sequence 11, Appl
42	73.2	2.6	2468	1	US-08-694-501-11	Sequence 11, Appl
43	72.6	2.6	1893	5	5438126-1	Patent No. 5438126
44	68.6	2.4	1677	2	US-08-372-652-14	Sequence 14, Appl
45	68.6	2.4	1677	4	PCT-US95-16311-14	Sequence 14, Appl

ALIGNMENTS

RESULT 1
US-07-737-736B-6
Sequence 6, Application US/0737736B
Patent No. 5260199
GENERAL INFORMATION:
APPLICANT: Deluca, Hector F.
APPLICANT: Ross, Troy K.
APPLICANT: Prahli, Jean M.
TITLE OF INVENTION: Method Of Producing
TITLE OF INVENTION: 1,25-Dihydroxyvitamin D3 Receptor Protein
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Carl R. Schwartz, Esq., c/o Quarles & Brady
STREET: 411 East Wisconsin Avenue
CITY: Milwaukee
STATE: Wisconsin
COUNTRY: U.S.A.
ZIP: 53202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/737,736B
FILING DATE: 19910730
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Schwartz, Carl R.
REGISTRATION NUMBER: 29,437
REFERENCE/DOCKET NUMBER: 96-296-2185-2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 414-277-5715
TELEFAX: 414-277-5774
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 2043 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: Double
MOLECULE TYPE: Linear
TOPOLOGY: Linear
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Rat
PUBLICATION INFORMATION:
AUTHORS: Burnester, James K.
AUTHORS: Wiese, Russell J.
AUTHORS: Maeda, No. 5260199uyo
AUTHORS: Deluca, Hector F.
TITLE: Structure and regulation of the rat

TITLE: 1,25-dihydroxyvitamin D3 receptor
JOURNAL: Proc. Natl. Acad. Sci. U.S.A.
VOLUME: 85
PAGES: 9499-9502
DATE: December-1988
US-07-737-736B-6

Query Match 10.0%; Score 278.8; DB 1; Length 2043;
Best Local Similarity 55.6%; Pred. No. 4e-51;
Matches 629; Conservative 0; Mismatches 482; Indels 21; Gaps 4;

288 ccccaaatcctcgtatgtgtgggacaagccactgctacatcattatgtatgaca 347
Db 155 ccccgactctgtgagtgctgtgagacgacccacacgcttccactcaatgatatgacc 214
Qy 348 tttgaagatcgaagggcttttcaagggcgccatgaacgaccccgctgagttgc 407
Db 215 tttgaagctcgaagggcttttcaagggcgccatgaacgaccccgctgagttgc 274
Qy 408 cccctccgaagggcgctgtgagatcaccccggaagaccccgagctgagccagctgc 467
Db 275 cccctcc--aattgagattgcccgcacatccacgaagacacccgcacactgcccagcctgc 331
Qy 468 cgcctgcaagctgctgtgagagcgagcaatgaagaagagatgatcatgtccgacagagcc 527
Db 332 cggctcgaagcctgtgtgagatcgccgatgtaagaggttccatctgacagatgagag 391
Qy 528 gtgagagagagggcgctgtatgaagcggagaaagatgaagcgaagagagctaccca 587
Db 392 gtacagcctgaagagagagatgattatgaagagaaagagagagccttgaagagagat 451
Qy 588 ctggagatgcaagggcgctgtgagagagagcagagatgatcatgaagggcgctgagcgt 647
Db 452 ctgag---gcccagctatctgaagaaacacagacacatcattccgctgagagcc 508
Qy 648 cagatgaagaacttgaactacatctccatctcaagaattccgctgagcag----- 702
Db 509 caccacagagacctgtgagacccacacacacacacacacacacacacacacacacacac 568
Qy 703 -gggtgttagcagtgctgtgagtgctgtgagtgctgtgagtgctgtgagtgctgtgag 761
Db 569 atgacag 628
Qy 762 gctgcaagtgagagcagtgctgtgagagagagagagagagagagagagagagagagag 821
Db 629 tcttctcagagctgtgacgtgacacacacacacacacacacacacacacacacacacac 668
Qy 822 cgggggag 861
Db 689 tccacactgagatgag 748
Qy 882 attctccctgtgctgac 941
Db 749 ccttctcagagctgtgac 808
Qy 942 agcttgcgaagtgctgtgag 1001
Db 809 ggccttgcgaagtgctgtgag 868
Qy 1002 ctgaagggggcgctgtgag 1061
Db 869 ctttaagtgcaagccac 928
Qy 1062 actggaacctgagagtgctgtgagagagagagagagagagagagagagagagagagag 1112
Db 929 gatattgcttgcgag 988
Qy 1113 ggtgtgcttgcgaagtgctgtgagagagagagagagagagagagagagagagagagagag 1172
Db 989 gctgggagac 1048
Qy 1173 ctgcagctgcatgag 1232

Db 1049 ctgaacttcaatgag 1108
Qy 1233 cggccagtgctgtgag 1292
Db 1109 cggccagtgctgtgag 1168
Qy 1293 ctgaacttcaatgag 1352
Db 1169 ctgaacttcaatgag 1228
Qy 1353 atcaatgctatgctcag 1404
Db 1229 atgacgaag 1280

RESULT 2
US-07-737-736B-5
Sequence 5, Application US/0773736B
Patent No. 5260199.

GENERAL INFORMATION:

APPLICANT: Deluca, Hector F.
APPLICANT: Ross, Troy K.
TITLE OF INVENTION: Method Of Producing
TITLE OF INVENTION: 1,25-dihydroxyvitamin D3 Receptor Protein
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:

ADDRESS: Carl R. Schwartz, Esq., c/o Quarles & Brady
STREET: 411 East Wisconsin Avenue
CITY: Milwaukee
STATE: Wisconsin
COUNTRY: U.S.A.
ZIP: 53202

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/737,736B
FILING DATE: 19910730

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Schwartz, Carl R.

REGISTRATION NUMBER: 29,437

REFERENCE/DOCKET NUMBER: 96-296-2185-2

TELECOMMUNICATION INFORMATION:

TELEPHONE: 414-277-5715

TELEFAX: 414-277-5774

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:

LENGTH: 1399 base pairs

TYPE: NUCLEIC ACID

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: cDNA

HYPOTHETICAL: NO

ANTI-SENSE: NO

ORIGINAL SOURCE:

ORGANISM: Homo sapiens

PUBLICATION INFORMATION:

AUTHORS: Baker, Andrew R.

AUTHORS: McDonnell, Donald P.

AUTHORS: Hughes, Mark

AUTHORS: Crisp, Tracey M.

AUTHORS: Mangelsdorf, David J.

AUTHORS: Haussler, Mark R.

AUTHORS: Pike, J. W.

AUTHORS: Shine, John

AUTHORS: O'Malley, Bert W.

TITLE: Cloning and expression of full-length cDNA

TITLE: encoding human vitamin D receptor

JOURNAL: Proc. Natl. Acad. Sci. U.S.A.
 VOLUME: 85
 PAGES: 3294-3298
 DATE: May-1988
 US-07-737-736B-5

Query Match 9.7% Score 272.8; DB 1; Length 1399;
 Best Local Similarity 55.1%; Pred. No. 6.8e-50;
 Matches 638; Conservative 0; Mismatches 487; Indels 33; Gaps 4;

288 ccccaatcgcgctgtagtgggagcaagccactggtatcactcaatgcatgaca 347
 176 ccccgagctgtaggggtgtaggagccgagccactggtcacttcaatgcatgacc 235
 348 tttgagagtgtaggggttcttcaggagggcgagaaagcgagcccggtgagttgc 407
 236 tctgagagctgtaggggttcttcaggagggcgagaaagcgagcccggtgagttgc 295
 408 cccctcagagggcgagctgtaggagccgagaaagcccgagctgtaggagctgagc 467
 296 cctctc---aacggagagctgtaggagccgagaaagcccgagctgtaggagctgagc 352
 468 cgcctgtaggagctgtaggagggcgagctgtaggagaaagagatgcatgctgagagggcc 527
 353 cggctcgaagcgtgtaggagctgtaggagaaagagatgcatgctgagagggcc 412
 528 gggagagagagggcgagctgtaggagaaagagatgcatgctgagagggcc 587
 413 gtcgagagagagggcgagctgtaggagaaagagatgcatgctgagagggcc 472
 588 ctggagagagagggcgagctgtaggagaaagagatgcatgctgagagggcc 647
 473 ctgagc---gcccagagctgtaggagaaagagatgcatgctgagagggcc 529
 648 cagagtagaagagctgtaggagaaagagatgcatgctgagagggcc 707
 530 caccatagagagctgtaggagaaagagatgcatgctgagagggcc 589
 708 ctgagagagagggcgagctgtaggagaaagagatgcatgctgagagggcc 749
 590 gtcgagagagagggcgagctgtaggagaaagagatgcatgctgagagggcc 649
 750 tcgagagagagggcgagctgtaggagaaagagatgcatgctgagagggcc 809
 650 tctggagagagggcgagctgtaggagaaagagatgcatgctgagagggcc 709
 810 tctctgagagggcgagctgtaggagaaagagatgcatgctgagagggcc 869
 710 tctctgagagggcgagctgtaggagaaagagatgcatgctgagagggcc 769
 870 ggggagagagggcgagctgtaggagaaagagatgcatgctgagagggcc 929
 770 ctgagagagagggcgagctgtaggagaaagagatgcatgctgagagggcc 829
 930 aagagagagagggcgagctgtaggagaaagagatgcatgctgagagggcc 989
 830 caaagagagagggcgagctgtaggagaaagagatgcatgctgagagggcc 889
 990 cagagagagagggcgagctgtaggagaaagagatgcatgctgagagggcc 1049
 890 cagagagagagggcgagctgtaggagaaagagatgcatgctgagagggcc 949
 1050 ttcagagagagggcgagctgtaggagaaagagatgcatgctgagagggcc 1100
 950 ttcagagagagggcgagctgtaggagaaagagatgcatgctgagagggcc 1009
 1101 gaaagagagagggcgagctgtaggagaaagagatgcatgctgagagggcc 1160
 1010 gacagagagagggcgagctgtaggagaaagagatgcatgctgagagggcc 1069
 1161 atgagagagagggcgagctgtaggagaaagagatgcatgctgagagggcc 1220

Db 1070 GAGCTGAGAGAGCTGAGTGCATGAGAGAGAGATGCTCTGCTATGCTGATC 1129
 Qy 1221 tctcccaagagggcgagctgtaggagaaagagatgcatgctgagagggcc 1280
 Db 1130 gtcctcccaagagggcgagctgtaggagaaagagatgcatgctgagagggcc 1189
 Qy 1281 ttcgagagagggcgagctgtaggagaaagagatgcatgctgagagggcc 1340
 Db 1190 ctcgagagagggcgagctgtaggagaaagagatgcatgctgagagggcc 1249
 Qy 1341 ttcgagagagggcgagctgtaggagaaagagatgcatgctgagagggcc 1400
 Db 1250 ctcgagagagggcgagctgtaggagaaagagatgcatgctgagagggcc 1309
 Qy 1401 cagagagagggcgagctgtaggagaaagagatgcatgctgagagggcc 1418
 Db 1310 aagagagagggcgagctgtaggagaaagagatgcatgctgagagggcc 1327

RESULT 3
 US-08-459-489-1
 : Sequence 1, Application US/08459489
 : Patent No. 568574
 : GENERAL INFORMATION:
 : APPLICANT: David D. Moore et al.
 : TITLE OF INVENTION: CAR RECEPTORS AND RELATED
 : TITLE OF INVENTION: MOLECULES AND METHODS
 : NUMBER OF SEQUENCES: 10
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Fish & Richardson
 : STREET: 225 Franklin Street
 : CITY: Boston
 : STATE: Massachusetts
 : COUNTRY: U.S.A.
 : ZIP: 02110-2804
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
 : COMPUTER: IBM PS/2 Model 502 or 555X
 : OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)
 : SOFTWARE: WordPerfect (Version 5.0)
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/08/459,489
 : FILING DATE:
 : CLASSIFICATION: 435
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: 07/843,350
 : FILING DATE: February 26, 1992
 : ATTORNEY/AGENT INFORMATION:
 : NAME: Paul T. Clark
 : REGISTRATION NUMBER: 30,162
 : REFERENCE/DOCKET NUMBER: 00786/126001
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: (617) 542-5070
 : TELEFAX: (617) 542-8906
 : TELEX: 200154
 : INFORMATION FOR SEQ ID NO: 1:
 : SEQUENCE CHARACTERISTICS:
 : LENGTH: 1450
 : TYPE: nucleic acid
 : STRANDEDNESS: single
 : TOPOLOGY: linear
 : US-08-459-489-1

Query Match 6.1%; Score 170.6; DB 1; Length 1450;
 Best Local Similarity 57.5%; Pred. No. 5.8e-28;
 Matches 327; Conservative 0; Mismatches 239; Indels 3; Gaps 1;

Qy 889 cctctgagagggcgagctgtaggagaaagagatgcatgctgagagggcc 948
 Db 739 cctctgagagggcgagctgtaggagaaagagatgcatgctgagagggcc 798
 Qy 949 ccaagagagagggcgagctgtaggagaaagagatgcatgctgagagggcc 1008

```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 1450
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-458-686-1

Query Match      6.1%; Score 170.6; DB 1; Length 1450;
Best Local Similarity 57.5%; Pred. No. 5.8e-28;
Matches 327; Conservative 0; Mismatches 239; Indels 3; Gaps 1;

Db 799 CTAAGGACCTGCCGCTCTCCCTCCCTGAGGACCAATGATCCCTTCCAGG 858
Qy 1009 gggcgcttgagctgtgtcaactgagattcaacacagtggtcaacggagactgaa 1068
Db 859 GAGCAGCTGTGGAATCTGTACATCGTACTCAATACCACTTCTGTCCAAACACAA 918
Qy 1069 cctggagatgtgagcgctgtctactgtcttggaagacatgaggt--ggtccacag 1125
Db 919 ACTTCTCTGCGGGCTCTGCTGCTACACATTTGAAGATGAGACCGCTGTGGGTTCCAG 978
Qy 1126 aacttctactgagcccatgtctgaattccactacatgtctgaaagactgagctgcatg 1185
Db 979 TAGAGTTTGGAGTTGCTCTTTCACTTCCATGAACTACGAACTGCAAGAGCTCCAG 1038
Qy 1186 aggaagatgtgtgtgtgtgagccatccctcttccccagacggccaggtgtgc 1245
Db 1039 AGCTGAGATGTGCTCTTGGCTGCTGCAATGGCCCTGTTCTCTGCAAGCTGGAATTA 1098
Qy 1246 tgcagacccgctgtgtgagccagctgacgaggaacattgcacattactctgaagctcaca 1305
Db 1099 CCCAGAGAGATGAGATTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1158
Qy 1306 ttgaatgcaatcgcccgagcctgtctataggttctgttccctgaagatcatggtatgc 1365
Db 1159 TCAAGGGCCAGCAGGAGGAGGCCCGGAGTCGATGATGATGATGATGATGATGATGATGAT 1218
Qy 1366 tcaccgagctgcgagcaatcaatgtctgagcacaccagcgctgtctgagctcagagaca 1425
Db 1219 TGCTGAGCTCGGAGCATTTATGAGGCTTACGGGCTACCAATTCAGCATCCAGGAGCC 1278
Qy 1426 tacacccttgtctagccctcatgag 1454
Db 1279 TGCTGAGCTGAGTGGCTGCTGCTCCAGGAG 1307

RESULT 4
US-08-458-686-1
; Sequence 1, Application US/08458686
; Patent No. 5710017
; GENERAL INFORMATION:
; APPLICANT: David D. Moore et al.
; TITLE OF INVENTION: CAR RECEPTORS AND RELATED
; TITLE OF INVENTION: MOLECULES AND METHODS
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)
; SOFTWARE: WordPerfect (Version 5.0)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/458,686
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/843,350
; FILING DATE: February 26, 1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Paul T. Clark
; REGISTRATION NUMBER: 30,162
; REFERENCE/DOCKET NUMBER: 00786/126001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 1:

```

```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 1450
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-458-686-1

Query Match      6.1%; Score 170.6; DB 1; Length 1450;
Best Local Similarity 57.5%; Pred. No. 5.8e-28;
Matches 327; Conservative 0; Mismatches 239; Indels 3; Gaps 1;

Db 889 cctgtgcgcccacatgagctgtgacatgtcaacctactgttcaagagcatcagacttg 948
Db 739 CTCTGCTACACACTTCCGAGACATACATTTCTATGCTAGTCAAGTACATCAAGTTTA 798
Qy 949 ccaaatcatctctactactcaaggaactgccatcgagagacagatctccctgtcgaag 1008
Db 799 CTAAGGACCTGCCGCTCTCCGTTCCGCTCCCATTTGAAGACCAAGATCCCTTCCAGG 858
Qy 1009 gggcgcttgagctgtgtcaactgagattcaacacagtggtcaacggagactgaa 1068
Db 859 GAGCAGCTGTGGAATCTGTACATCGTACTCAATACCACTTCTGTCTCCAAACACAA 918
Qy 1069 cctggagatgtgagcgctgtctactgtcttggaagacatgaggt--ggtccacag 1125
Db 919 ACTTCTCTGCGGGCTCTGCTGCTACACATTTGAAGATGAGACCGCTGTGGGTTCCAG 978
Qy 1126 aacttctactgagcccatgtctgaattccactacatgtctgaaagactgagctgcatg 1185
Db 979 TAGAGTTTGGAGTTGCTCTTTCACTTCCATGAACTACGAACTGCAAGAGCTCCAG 1038
Qy 1186 aggaagatgtgtgtgtgtgagccatccctcttccccagacggccaggtgtgc 1245
Db 1039 AGCTGAGATGTGCTCTTGGCTGCTGCAATGGCCCTGTTCTCTGCAAGCTGGAATTA 1098
Qy 1246 tgcagacccgctgtgtgagccagctgacgaggaacattgcacattactctgaagctcaca 1305
Db 1099 CCCAGAGAGATGAGATTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1158
Qy 1306 ttgaatgcaatcgcccgagcctgtctataggttctgttccctgaagatcatggtatgc 1365
Db 1159 TCAAGGGCCAGCAGGAGGAGGCCCGGAGTCGATGATGATGATGATGATGATGATGATGAT 1218
Qy 1366 tcaccgagctgcgagcaatcaatgtctgagcacaccagcgctgtctgagctcagagaca 1425
Db 1219 TGCTGAGCTCGGAGCATTTATGAGGCTTACGGGCTACCAATTCAGCATCCAGGAGCC 1278
Qy 1426 tacacccttgtctagccctcatgag 1454
Db 1279 TGCTGAGCTGAGTGGCTGCTGCTCCAGGAG 1307

RESULT 5
US-07-843-350C-1
; Sequence 1, Application US/07843350C
; Patent No. 5756448
; GENERAL INFORMATION:
; APPLICANT: David D. Moore et al.
; TITLE OF INVENTION: CAR RECEPTORS AND RELATED
; TITLE OF INVENTION: MOLECULES AND METHODS
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)

```


QY 405 tgccttcgga

Dh	434	TATGGCTGGCCBGGGTGGGGAAACCTGCCAGATGAGACCTTTTATGGGGCCAAAGTGCAG	493
OY	462	gacctgcgcctcgcgaagaatgctccgtgagaagcgcatgaaagaagatgatcatgtccgac	521
Dh	494	CAGTGGCCGGCTCCGAACTGCAAGAGGACAGGGATGAGGAGAGAGTCCGCTCTTTTGAA	553
OY	522	gaagcccgctgagagagagcgcgcttgatacaagcggagaagaagtgaacggaacgagact	581
Dh	554	GAACAGATCCGAGAAAGAAAGATTCCGAAACAGCAGCAGCAGCAGTCCACAGTCCAGTGC	613
OY	582	cagccaactcggagagtgctcaggggtctaacagaagagcagcggatgatgatcaaggagctgatg	641
Dh	614	CAGTCACTGTGGGGCCG-----CAGGGCAGCAGCAGCTCAGCCTCTGGGCTTGGGGC	666
OY	642	gacgctcagaagaanaaccttgacataacctctccattcaagaattccgctgcgca	701
Dh	667	TTCCCCCTGGTGA-----TTT	682
OY	702	ggagtgcttaacgtagtgcgtcgagattgccaaggtctcgcagcggcccatgagaagaga	761
Dh	683	GAGCGACGACACCCAGGGGCTCCGGGGAAGAGAGGGTGTCCAGCTTAACAGGGCTCAAGA	742
OY	762	gctgcgaaggtgagccaggtccggaaagatctgtgctcttgaa--ggtctctgcagc	819
Dh	743	CTAATGATCCAGAGTTGGTGGGGCCCACTGCAAGTCAACAAAGCTCTTCTCCGAC	802
OY	820	tgcgggggagatgagtgcaagtctcgcgaactacaacaccccgccgcgacagtgcgggaaag	879
Dh	803	CAGCCCAAAAGTCAGCCCTGGCCCTCGGGGCGAGACCCCAAGTCCGGAATGCGCCGACG	862
OY	880	agatctctccctcgtcgcacacatgctgacatgltcaactacatacgttcaagagcatca	939
Dh	863	CAA-----CGGTTTGGCCCACTTCACGGAGGTGGCAATCATCTCAGTCCAGAGGATGG	914
OY	940	tcaagcttgccaagaatcatctccactctcaaggagacttgcacatcgagagacagactctcc	999
Dh	915	TGGACTTGGCTTAACCAAGTGCCTGTGTTTCTGCAAGTGGGCCGGGAGGACCAAGTGC	974
OY	1000	tgtctgaaggggacgcttgcggagctgtgtcaacttgaaatccaacacagtglttcaacgcg	1055
Dh	975	TCTCGAAGGCAATCCACTATGAGATCATGCTGCTAAGACAGCAGCAGCGCTCAACACAG	1034
OY	1060	agactcgg-----acctggagtggtgcggcgctgtccactcgtcttggaaagcactgag	1111
Dh	1035	AGACAGAGTGTATCACTCTTTGAAGACTTCACCTACAGCAAGACGACGCTTCCACCGTG	1094
OY	1114	gtgctctccacaaactctactctgagcccaatgctgtaattccaactacatacgtcgtgaagac	1173
Dh	1095	CAGGCGTCGACAGGTGGAATTCATCAACCCATCTTCAGATTCTCGGGGGCAGTGGCGGGC	1155
OY	1174	tgcacgtcgcacatgagagagatgctgtctatgcaagcaatctccctctctccccaagc	1233
Dh	1155	TGGGCTGGAGAGCGGTAATGAGCCCTGCTATCCCATCAAAATCTTCTGGCGGACCT	1214
OY	1234	gcccaggtgtctcgcacaacgcgcgtgtgttgccaagctgcaaggagaattcgcacatactc	1293
Dh	1215	GGCTTAATGTGACAGAGCCGGGGCGGTGGAGGCGTTTCACAGACCCCTACGTGAGGCGCC	1274
OY	1294	tgaagtcctcaatgtaatgcaatcggccccagcgtgctcatagttgtctgttctctgaaga	1353
Dh	1275	TGCTGTCTCTACAGCGCGTTACAAAGGCGCCAGACAGCTGTGGCTTCCCGGCAATGTCA	1344
OY	1354	tcatggtctatgctcacaocgagctcgcgagataaagtgtagaacacccaagcgctgcgc	1413
Dh	1335	TGAAGCTGGTGAAG-----CTGCGCACTTGAAGCTGTGTGACATCGGAGCAGGGCTTTGG	1388
OY	1414	gcatacgaagacatacacccttgcgtacagccctcatcagcagaagtggttctgcgaatcacag	1473
Dh	1389	CTTTGGGGGCTTCAGAGCAAGAAAGCTGGCGGCTGTGTGCGAGATCTGGGACGTGCAGC	1448
OY	1474	gttagctgagcgctgcctcttggtgacacc	1503
Dh	1449	AGTGAAGGGCTTGGCCACCCAGCCCCACAC	1478


```

1 TELEPHONE: (908) 594-4283
2 TELEFAX: (908) 594-4720
3
4 INFORMATION FOR SEQ ID NO: 1:
5
6 SEQUENCE CHARACTERISTICS:
7
8 LENGTH: 2030 base pairs
9
10 TYPE: nucleic acid
11
12 STRANDEDNESS: single
13
14 TOPOLOGY: linear
15
16 MOLECULE TYPE: DNA (genomic)
17
18 HYPOTHEetical: NO
19
20 ANTI-SENSE: NO
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
102
```

Query Match	3.58;	Score 97.6;	DB 2;	Length 2030;
Best Local Similarity	54.88;	Pred. No. 3e-12;		
Matches 215; Conservative	0;	Mismatches 174;	Indels 3;	Gaps 1.

[illegible]

RESULT 12
PCT-US95-13924-1
: Sequence 1, Application PC/TUS9513924
GENERAL INFORMATION:
APPLICANT: Friedman, Eitan
APPLICANT: Holloway, M. Katharine
APPLICANT: Rodon, Gideon
APPLICANT: Rutledge, Su Jane
APPLICANT: Schmidt, Azriel
APPLICANT: Vogel, Robert
TITLE OF INVENTION: METHOD FOR FINDING RECEPTOR POTENTIATORS
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Merck & Co., Inc.
STREET: 126 East Lincoln Avenue
CITY: Rahway
STATE: New Jersey
COUNTRY: US
ZIP: 07065-0907
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA: PCT/US95/13924

```

1      FILING DATE:
2      CLASSIFICATION:
3      ATTORNEY/AGENT INFORMATION:
4      NAME:  Ouellette, Carol S.
5      REGISTRATION NUMBER: 35,330
6      REFERENCE/DOCKET NUMBER: 19327 PCT
7      TELECOMMUNICATION INFORMATION:
8      TELEPHONE: (908) 594-3809
9      TELEFAX: (908) 594-4720
10     INFORMATION FOR SEQ ID NO: 1:
11     SEQUENCE CHARACTERISTICS:
12     LENGTH: 2030 base pairs
13     TYPE: nucleic acid
14     STRANDEDNESS: single
15     TOPOLOGY: linear
16     MOLECULE TYPE: DNA (genomic)
17     HYDROTHERMAL: NO
18     ANTI-SENSE: NO
19     PCT-US95-13924-1

```

Query Match	3.58;	Score 97.6;	DB 4;	Length 2030;
Best Local Similarity	54.88;	Pred. No. 3e-12;		
Matches 215; Conservative	0;	Mismatches 174;	Indels 3;	Gaps 1.

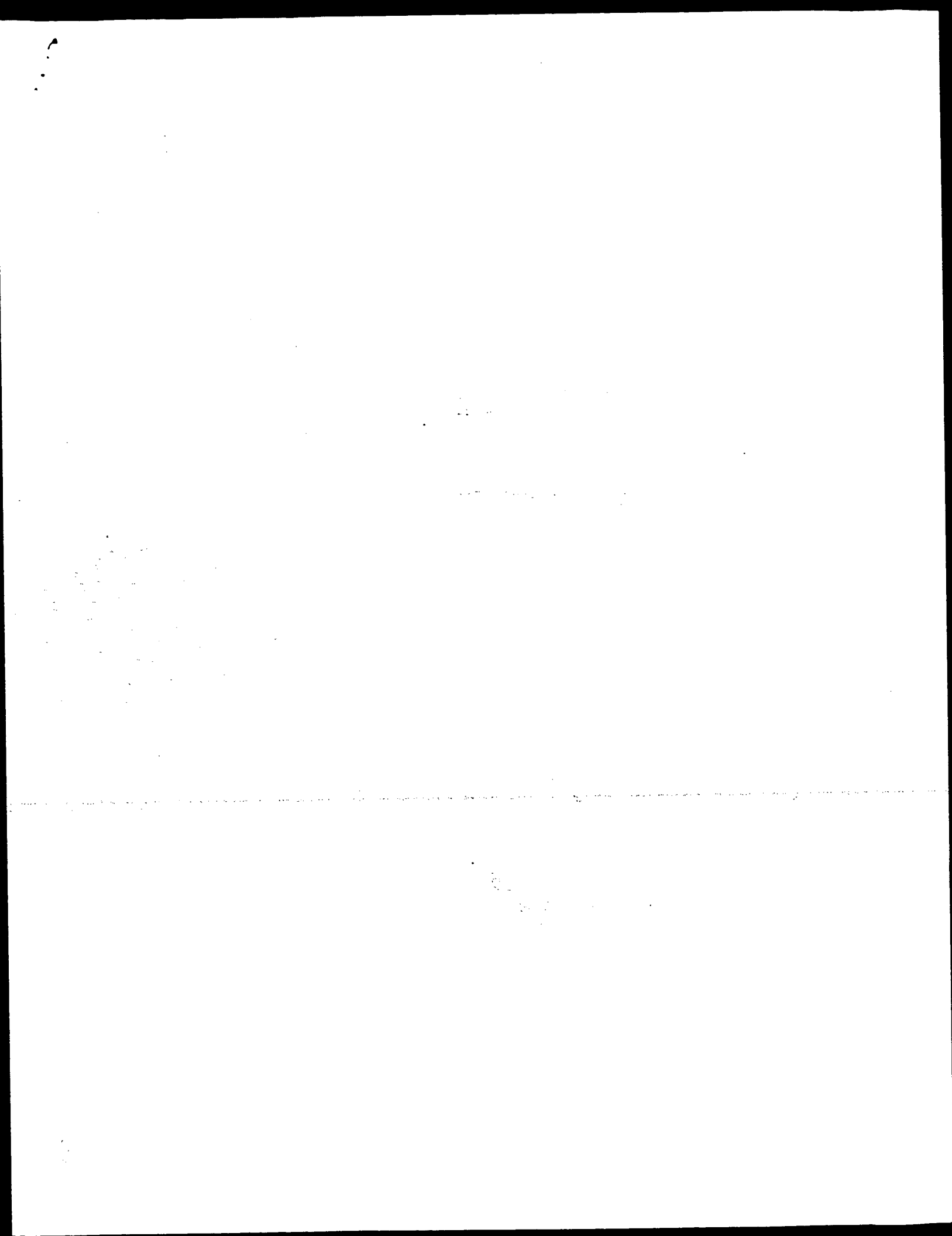
Oy 211 ctgacttgatcactgtggaagacacaaagctgtctcctgtaaaagcccagatgtcaacgcag 270
 Db 417 CAGATGGGATATCCAGATCCGACAAAGGAAACCAAGGCCAAGCCAAAGAGGCCAG 476
 Oy 271 atgagaagtctggaaggtcccaaatctgcgcgtgatgtgtgggaacaagccactgtgctatc 330
 Db 477 CCCGGAATATCTGGGCGACAGAGTTGGCGGTGTCTGTGGGGACAAGGCTCCGGCTTCC 536
 Oy 331 actcaatgtcatatgacatgtgaagatgaagatgaagaggtcttttccaaagagggccatgaaacga 380
 Db 537 ACTAACAAGGTCTAGCTAGCTGCACAAAGCTGCAGAGGCTCTTCCGGCGCAGATGTGCTCGTG 586
 Oy 391 aagcccgctgaaggtgtgcacctctcggaagaagcg---ccctgagatcaaccggaagaagccc 447
 Db 597 GTGGGGCCAGCGCGCTATGTCTCTGCCCCGGGTGGCGGAACCTGCCAGATGACGCTTTCAATGC 656
 Oy 448 ggcgaacagttgcacagccttgcgcgcctctgcgcaagttgctcttgaaagacggtcatgaagaaga 507
 Db 657 GCGCGAATGGCAGAGAGTGGCGGTGGCGCAAGTGCACAAAGAGGCGGATGAGGGAGCAGT 716
 Oy 508 tgatcatgtctcgacagagcgcttgaggaagagcggtgcaccttgatcaagcggagaagaagtgt 567
 Db 717 GCGTCTCTTTGTAAGAACAGATCCGGGAAGAAGATTGCGMAACACAGCAGCAGAGAT 776
 Oy 568 aacggacagggacatcagccacttgggaagtgcag 599
 Db 777 CACAATCACAATGCCAGTCACTCTGTGGGGCGG 808

RESULT 13
PCF-US95-13931-1
Sequence 1, Application PC/TUS9513931
GENERAL INFORMATION:
APPLICANT: Friedman, Eitan
APPLICANT: Holloway, M. Katharine
APPLICANT: Rodan, Gideon
APPLICANT: Schmidt, Ariel
APPLICANT: Vogel, Robert
TITLE OF INVENTION: USE OF RECEPTOR POTENTIATORS
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Merck & Co., Inc.
STREET: 126 East Lincoln Avenue
CITY: Rahway
STATE: New Jersey
COUNTRY: US
ZIP: 07065-0907
COMPUTER READABLE FORM:


```

1 TITLE OF INVENTION: TREATMENT OF ACUTE PROMYELOCYTIC LEUKEMIA (APL)
2
3 NUMBER OF SEQUENCES: 11
4
5 CORRESPONDENCE ADDRESS:
6
7 ADDRESSEE: COOPER & DUNHAM LLP
8
9 STREET: 1185 AVENUE OF THE AMERICAS
10
11 CITY: NEW YORK
12
13 STATE: NEW YORK
14
15 COUNTRY: U.S.A.
16
17 ZIP: 10036
18
19 COMPUTER READABLE FORM:
20
21 MEDIUM TYPE: Floppy disk
22
23 COMPUTER: IBM PC compatible
24
25 OPERATING SYSTEM: PC-DOS/MS-DOS
26
27 SOFTWARE: PatentIn Release #1.24
28
29 CURRENT APPLICATION DATA:
30
31 APPLICATION NUMBER: US/08/095,728B
32
33 FILING DATE: 21-JUL-1993
34
35 CLASSIFICATION: 435
36
37 PRIOR APPLICATION DATA:
38
39 APPLICATION NUMBER: US 07/673,838
40
41 FILING DATE: 22-MAR-1991
42
43 ATTORNEY/AGENT INFORMATION:
44
45 NAME: WHITE, JOHN P
46
47 REGISTRATION NUMBER: 28,678
48
49 REFERENCE/DOCKET NUMBER: 38694-A
50
51 TELECOMMUNICATION INFORMATION:
52
53 TELEPHONE: (212) 278-0400
54
55 TELEFAX: (212) 391-0525
56
57 INFORMATION FOR SEQ ID NO: 3:
58
59 SEQUENCE CHARACTERISTICS:
60
61 LENGTH: 2928 base pairs
62
63 TYPE: nucleic acid
64
65 STRANDEDNESS: single
66
67 TOPOLOGY: linear
68
69 MOLECULE TYPE: cDNA
70
71 IMMEDIATE SOURCE:
72
73 CLONE: hRAR ALPHA
74
75 FEATURE:
76
77 NAME/KEY: CDS
78
79 LOCATION: 103..1488
80
81 OTHER INFORMATION:
82
83 US-08-095-728B-3

```



GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 18, 2001, 16:54:40 ; Search time 211.71 Seconds

(without alignments)
4971.928 Million cell updates/sec

Title: US-09-143-828-3

Perfect score: 2802
Sequence: 1 tgaattcgtggcctgctg.....aaaaaaaaaaaaaaaaaa 2802

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 480022 seqs, 187831343 residues

Total number of hits satisfying chosen parameters: 960044

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

N.Geneseq_36:*

- 1: /cgn2_2/gcgdata/geneseq/geneseqn/NA1980.DAT:*
- 2: /cgn2_2/gcgdata/geneseq/geneseqn/NA1981.DAT:*
- 3: /cgn2_2/gcgdata/geneseq/geneseqn/NA1982.DAT:*
- 4: /cgn2_2/gcgdata/geneseq/geneseqn/NA1983.DAT:*
- 5: /cgn2_2/gcgdata/geneseq/geneseqn/NA1984.DAT:*
- 6: /cgn2_2/gcgdata/geneseq/geneseqn/NA1985.DAT:*
- 7: /cgn2_2/gcgdata/geneseq/geneseqn/NA1986.DAT:*
- 8: /cgn2_2/gcgdata/geneseq/geneseqn/NA1987.DAT:*
- 9: /cgn2_2/gcgdata/geneseq/geneseqn/NA1988.DAT:*
- 10: /cgn2_2/gcgdata/geneseq/geneseqn/NA1989.DAT:*
- 11: /cgn2_2/gcgdata/geneseq/geneseqn/NA1990.DAT:*
- 12: /cgn2_2/gcgdata/geneseq/geneseqn/NA1991.DAT:*
- 13: /cgn2_2/gcgdata/geneseq/geneseqn/NA1992.DAT:*
- 14: /cgn2_2/gcgdata/geneseq/geneseqn/NA1993.DAT:*
- 15: /cgn2_2/gcgdata/geneseq/geneseqn/NA1994.DAT:*
- 16: /cgn2_2/gcgdata/geneseq/geneseqn/NA1995.DAT:*
- 17: /cgn2_2/gcgdata/geneseq/geneseqn/NA1996.DAT:*
- 18: /cgn2_2/gcgdata/geneseq/geneseqn/NA1997.DAT:*
- 19: /cgn2_2/gcgdata/geneseq/geneseqn/NA1998.DAT:*
- 20: /cgn2_2/gcgdata/geneseq/geneseqn/NA1999.DAT:*
- 21: /cgn2_2/gcgdata/geneseq/geneseqn/NA2000.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the total being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2802	100.0	2802	20 X56243	Human vitamin D re
2	2715.4	96.9	2850	20 X78808	Human nm7-1 CDNA.
3	2702	96.4	3243	20 X59974	DNA encoding an in
4	2640.2	94.2	2810	20 X56242	Human vitamin D re
5	2586.2	92.3	3057	20 X59975	DNA encoding an in
6	2586.2	92.3	3093	20 X78807	Human nm7 CDNA.
7	1866.6	66.6	2146	20 X207997	Human pregnane x r
8	1466.6	52.3	2068	20 X89080	Human steroid and
9	1422	50.7	1422	20 X59968	DNA encoding an in
10	1330.2	47.5	1374	20 X59967	DNA encoding an in
11	1305	46.6	1305	20 X59966	DNA encoding an in
12	425.2	15.2	644	20 X59972	SSO ID 12 of JP111

13	278.8	10.0	1960	19 V03129	CDNA encoding rat
14	278.8	10.0	2043	14 Q51425	Rat vitamin D rece
15	272.8	9.7	1382	20 X34789	Nucleotide sequenc
16	272.8	9.7	1463	20 X34788	Nucleotide sequenc
17	272.8	9.7	1574	20 X34793	Nucleotide sequenc
18	272.8	9.7	3510	20 Z07545	Human vitamin D re
19	272.8	9.7	4604	19 V41347	Human vitamin D re
20	271.2	9.2	1399	14 Q51424	Human vitamin D re
21	257	9.2	2191	17 T66499	Xenopus orphan rec
22	236	8.4	1071	20 X16596	Rat vitamin D rece
23	236	8.4	3382	19 V18518	CDNA encoding rat
24	236	8.4	3382	20 X16606	Rat vitamin D rece
25	215	7.7	1534	20 X34790	Nucleotide sequenc
26	188.6	6.7	1404	20 X16597	Human vitamin D re
27	185.4	6.6	1404	19 V03130	CDNA encoding rat
28	181.6	6.5	1280	20 X80215	Mouse nuclear rece
29	172.2	6.1	1450	18 T92305	Constitutively act
30	170.6	6.1	1450	14 Q46131	Human CAR receptor
31	170.6	6.1	1450	20 X33994	Human CAR receptor
32	168.6	6.0	1361	20 X24003	Mouse CAR receptor
33	108.4	3.9	468	20 X80217	Mouse nuclear rece
34	97.6	3.5	1898	18 T79634	DNA encoding human
35	97.6	3.5	1979	17 T27616	Human foetal lung
36	97.6	3.5	2030	15 Q53138	Human recombinant
37	97.6	3.5	2030	17 T18996	Human steroid rece
38	97.6	3.5	2030	17 T30031	NER receptor poten
39	97.2	3.5	1813	16 Q88760	Human ubiquitinous n
40	91.4	3.3	2928	13 Q29338	RAR-alpha gene. S
41	91.4	3.3	2928	20 V64991	Human RAR-alpha CD
42	91.4	3.3	2940	10 N80124	DNA of clone phRAR
43	91.4	3.3	3036	13 Q29334	myl/RAR-alpha fusl
44	91.4	3.3	3036	19 V20474	Human PML/RARalpha
45	91.4	3.3	3036	20 V64990	Human myl/RAR-alpha

ALIGNMENTS

RESULT 1

ID X56243 standard: CDNA: 2802 BP.

AC X56243;

DT 16-JUL-1999 (first entry)

XX

DE Human vitamin D receptor related gamma 2 protein encoding CDNA.

XX

KX Human: vitamin D receptor related protein; VDR; obesity; diabetes;

KW anorexia; rheumatoid arthritis; lipoprotein defect; hyperlipidaemia;

KW hypercholesterolaemia; hyperlipoproteinaemia; osteoporosis; tumour;

KW hyperproliferative skin disorder; hyperthyroidism; ss.

XX

OS Homo sapiens.

XX

PN WO9919354-A1.

XX

PD 22-Apr-1999.

XX

EF 31-AUG-1998; 98WO-SE01548.

XX

PR 31-MAR-1998; 98SE-0001148.

PR 14-OCT-1997; 97SE-0003745.

XX

PA (PHAA) PHARMACIA & UPJOHN AB.

XX

PI Berkenstam A, Dahlberg M.

XX

DR WPI: 1999-302508/25.

XX

PT P-PSDB: Y09516;

PT New vitamin D receptor related (VDR) polypeptides, useful for treating obesity, diabetes, anorexia and rheumatoid arthritis

Db	1741	tcgagagagagaaagtgcgccttcctttaaaaggccctggtcctggggaataccct	1800
Qy	1801	cagatcccaataaagtgtcacaagtgvtgnaagggacaaagcgacaaagataggccatctg	1860
Db	1801	cagatcccaataaagtgtcacaagtgvtgnaagggacaaagcgacaaagataggccatctg	1860
Qy	1861	gggtctatgcgccacataacccaagttgttcggttccttcgaagcttttaattgctaccccta	1920
Db	1861	gggtctatgcgccacataacccaagttgttcggttccttcgaagcttttaattgctaccccta	1920
Qy	1921	ataatctgcgtccacacttccacactgcgttcctccctccctccctccacagactctgtgggtc	1980
Db	1921	ataatctgcgtccacacttccacactgcgttcctccctccctccctccacagactctgtgggtc	1980
Qy	1981	aaaggctctatcatcatcgcaagtgcaatgaatactctgtgagtagtccctaaaggagaatga	2040
Db	1981	aaaggctctatcatcatcgcaagtgcaatgaatactctgtgagtagtccctaaaggagaatga	2040
Qy	2041	agccaaagagagcctgcgacacaaatgtccgaagaagcttgacagactctcatcttcggccacatcat	2100
Db	2041	agccaaagagagcctgcgacacaaatgtccgaagaagcttgacagactctcatcttcggccacatcat	2100
Qy	2101	tcgtgtgctctgtgataccaatttgaaacacttatataagacatgaataagtagtagcctgctgt	2160
Db	2101	tcgtgtgctctgtgataccaatttgaaacacttatataagacatgaataagtagtagcctgctgt	2160
Qy	2161	gggtgtacacagcaatctgaatcagaataaagaatccctgaagctcacagaagattatagttaaaana	2220
Db	2161	gggtgtacacagcaatctgaagctcagataatagatcctcgagctcacagaagattatagttaaaana	2220
Qy	2221	acaaacagaaacacacaaacaaatcttgatcacaaagagaaataataatgtgacaaagaacgac	2280
Db	2221	acaaacagaaacacacaaacaaatcttgatcacaaagagaaataataatgtgacaaagaacgac	2280
Qy	2281	acaaagagaaattccctcctgtgtgagatgtgcagactgtgaaatgcagacatgtggtaaccagaatga	2340
Db	2281	acaaagagaaattccctcctgtgtgagatgtgcagactgtgaaatgcagacatgtggtaaccagaatga	2340
Qy	2341	aggtctcccgagagaaatgaaatgactctgttaagagacaaagggacaaacaaatctgaagactgtgagctg	2400
Db	2341	aggtctcccgagagaaatgaaatgactctgttaagagagaaagggacaaacaaatctgaagactgtgagctg	2400
Qy	2401	tgtgtgatacttggtgtgtagagtgacgtctgtttgcacactgataaggggctggtgttctcccg	2460
Db	2401	tgtgtgatacttggtgtgtagagtgacgtctgtttgcacactgataaggggctggtgttctcccg	2460
Qy	2461	ggctcggaatgtgtgggtgtgtgtctctgtgacaaaggtaaagcgcgaacaaatcgaattaaacacacg	2520
Db	2461	ggctcggaatgtgtgggtgtgtgtctctgtgacaaaggtaaagcgcgaacaaatcgaattaaacacacg	2520
Qy	2521	gagaaagaaacaaatctaacatgcacacttaattctcgtgtaacacatctatcttcacaaagctaa	2580
Db	2521	gagaaagaaacaaacaaatctaacatgcacacttaattctcgtgtaacacatctatcttcacaaagctaa	2580
Qy	2581	gggtatgaaagtgccctgcgctcgtgtttatagccactctgtgagtaaaaaattttttgtgacattt	2640
Db	2581	gggtatgaaagtgccctgcgctcgtgtttatagccactctgtgagtaaaaaattttttgtgacattt	2640
Qy	2641	cacaaatctatctcttataataaggaattccacaaactgaagaactgttttggaagatgtgc	2700
Db	2641	cacaaatctatctcttataataaggaattccacaaactgaagaactgttttggaagatgtgc	2700
Qy	2701	cctgtggttaatgtcacaatcagaagcagaagaagaaattaaataatgttaactttgtgctaaana	2760
Db	2701	cctgtggttaatgtcacaatcagaagcagaagaagaaattaaataatgttaactttgtgctaaana	2760
Qy	2761	aaa 2802	
Db	2761	aaa 2802	
RESULT	2		

```

X78808
ID X78808 standard; cDNA; 2850 BP.
XX
AC X78808:
XX
DT 06-SEP-1999 (first entry)
XX
DE Human nNR7-1 cDNA.
XX
KM nNR7; nNR7-1; nuclear trans-acting receptor protein; human; regulator;
KW identification; downstream target gene; cell proliferation;
KM cell development; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers.
FT CDS 104..1525
FT /tag= a
FT /product= "nNR7-1"
XX
PN W09931129-A1.
XX
PD 24-JUN-1999.
XX
PF 11-DEC-1998; 98MO-US26364.
XX
PR 14-OCT-1998; 98US-0104251.
PR 12-DEC-1997; 97US-0069401.
XX
PA (MERI ) MERCK & CO INC.
XX
PI Chen F;
XX
DR WPI: 1999-405024/34.
DR P-PSDB; Y25411.
XX
PT DNA encoding human nuclear receptors nNR7 and nNR7-1
XX
PS Claim 28; Fig 4A-C; 80pp; English.
XX
XX
XX This invention describes the isolation of the novel human nuclear
XX receptors nNR7 or nNR7-1. The nNR7 and nNR7-1 proteins are useful in
XX the identification of downstream target genes and ligands regulating its
XX activity. The nuclear receptor is involved in the regulation of in vivo
XX cell proliferation and/or cell development. The nNR7 and nNR7-1
XX polynucleotides, expression vectors and host cells are useful for the
XX recombinant production of the protein.
XX
SQ Sequence 2850 BP; 691 A; 736 C; 787 G; 636 T; 0 other:

Query Match 96.9%; Score 2715.4; DB 20; Length 2850;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 2741; Conservative 0; Mismatches 6; Indels 2; Gaps 2

QY 9 tgggctctgtggtgtgtgtctgtgcagccccccctgagggccaagagacagcagcatgacagtc 68
DB 54 ttggccctgtggtgttagtctgtgcag-ccccctgagggccaagagacagcagcatgacagtc 112
QY 69 accaggaactaaccaacttcaaggaggggtccctcagagacacttcataccctgcacagt 128
DB 113 accaggaactaaccaacttcaaggaggggtccctcagagacacttcataccctgcacagt 172
QY 129 gctgcgctgtgagtgtgcttcaaaccttcaaggagggcaccagaacaaactggaagtga 188
DB 173 gctgcgctgtgagtgtgcttcaaaccttcaaggagggcaccagaacaaactggaagtga 232
QY 189 cccaagaagaagctggaacatgctgacctgtgtacactgtgaggaacacagatgtgttct 248
DB 233 cccaagaagaagctggaacatgctgacctgtgtacactgtgaggaacacagatgtgttct 292
QY 249 ggaagaccccggttcaacgagatgaggaagtcggaggtcccaaatctgcgtgatgt 308

```

Db	293	ggaagaagcccaagtgtcacacgcagatgagggaagtcgagaggtccccaatactgcgtgatatgt	352
QY	309	gggggcagaagggccactggtcattcaattcaatgtcatgcatggtgaagatgcaagggcttc	368
Db	353	gggggcagaagggccactggtcattcaattcaatgtcatgcatggtgaagatgcaagggcttc	412
QY	369	ttcaagagagggcactgaataaacgcgaagccgggtgtgaggtgtgccccttcgcgaaggggcgtctc	428
Db	413	ttcaagagagggcactgaataaacgcgaagccgggtgtgaggtgtgccccttcgcgaaggggcgtctc	472
QY	429	gagatcaccccggaagaaagaccggcggaacagtgcgaaggtctgcgcgcgtgcgaagtgctgtgaag	488
Db	473	ggagatcaccccggaagaaagaccggcggaacagtgcgaaggtctgcgcgcgtgcgaagtgctgtgaag	532
QY	489	aagcgcgcataagaaagaaatgatatcatatcttcggaacgaagacccgtgtggaagaaagagccggcttg	548
Db	533	aagcgcgcataagaaagaaatgatatcatatcttcggaacgaagacccgtgtggaagaaagagccggcttg	592
QY	549	atcaaaagcggaaagaaagatgtgaacacgggacagactgaagccaaactgtggagagtgcagggcgcaac	608
Db	593	atcaaaagcggaaagaaagatgtgaacacgggacagactgaagccaaactgtggagagtgcagggcgcaac	652
QY	609	gaaggagcagcggaatgatagtacagggagactgtgaatgcgtctaaagtataaataaccttggacact	668
Db	653	gaaggagcagcggaatgatagtacagggagactgtgaatgcgtctaaagtataaataaccttggacact	712
QY	669	aacctctcccaatttcacaagaatttcacgggtctgcgaaggggtgtgtcttaagctgtgcgtgcgaatttg	728
Db	713	aacctctcccaatttcacaagaatttcacgggtctgcgaaggggtgtgtgtcttaagctgtgcgtgcgaatttg	772
QY	729	ccaaaggtctctgcgcagggcccaatcgaagaggaagaaagctgtgcgaagtgtgaagccaaagtccggaata	788
Db	773	ccaaaggtctctgcgcagggcccaatcgaagaggaagaaagctgtgcgaagtgtgtgaagccaaagtccggaata	832
QY	789	gatactgtcctctcttggaaaggtgtctctctgcacgtctgcggggggggaagaaatggacagtctcttggaac	848
Db	833	gatactgtcctctcttggaaaggtgtctctctgcacgtctgcggggggggaagaaatggacagtctcttggaac	892
QY	849	tacaaaccccacagcgcgaacaggtgcggggaaagaaagactctctccctgtctgcacacatgtggt	908
Db	893	tacaaaccccacagcgcgaacaggtgcggggaaagaaagactctctccctgtctgcacacatgtggt	952
QY	909	gacactgtcaaacctctacgtttccaaagagatatataagctttggccaaagatactctccatacttc	968
Db	953	gacactgtcaaacctctacgtttccaaagagatatataagctttggccaaagatactctccatacttc	1012
QY	969	aaggagactgtcccaatcgcgaagacagatctccctcgtctgaaggggggcgccttcctgaagctgtgt	1028
Db	1013	aaggagactgtcccaatcgcgaagacagatctccctcgtctgaaggggggcgccttcctgaagctgtgt	1072
QY	1029	caaatgtgattccaacacaaagtggttcacaaagcgggaacacgcggaacctgaggagtgtgcgcggctgt	1088
Db	1073	caaatgtgattccaacacaaagtggttcacaaagcgggaacacgcggaacctgaggagtgtgtgcgcggctgt	1132
QY	1089	tcctcacactgtctggaaagacacactgcgaaggtgtcttcacaaacaaactcttaactgtgaagcccaatgtctg	1148
Db	1133	tcctcacactgtctggaaagacacactgcgaaggtgtctcttcacaaacaaactcttaactgtgaagcccaatgtctg	1192
QY	1149	aaatctcacactaactgtctgaagaaagctgtgcagctgtcatgtgaagagagagatagtgtctgtatgcag	1208
Db	1193	aaatctcacactaactgtctgaagaaagctgtgcagctgtcatgtgaagagagagatagtgtctgtatgcag	1252
QY	1209	ggcaactccacctctcttcgcccaagacccgcgaaggtgtctctgcagaacacggcggtgtgtgaacag	1266
Db	1253	ggcaactccacctctcttcgcccaagacccgcgaaggtgtctctgcagaacacggcggtgtgtgtgaacag	1312
QY	1269	ctggcagagagcaattcgcacattactctgaaggtctctaaattgtgaatgtgaatgtgcacacgcct	1328
Db	1313	ctggcagagagcaattcgcacattactctgaaggtctctaaattgtgaatgtgaatgtgcacacgcct	1372
QY	1329	ggtcatatggtttcttcttgcgcgaagaatcataggtctatgcacgaaggtctccgcgaatcaat	1388
Db	1373	ggtcatatggtttcttcttgcgcgaagaatcataggtctatgcacgaaggtctccgcgaatcaat	1432

QY	1389	gccaagacacacacacagcgcgctgcgtcgtacatccaaagaaataacacccctttgtctacgcccctc	1448
Db	1433	gtctccagccacacacacagcgcgctgcgtcgtacatccaaagaaataacacccctttgtctacgcccctc	1492
QY	1449	atcgacaggtgtctgttcctccgcatcaacagctgcagctctgagcgcgctgcgccccttgggtgcacacccctcga	1508
Db	1493	atcgacaggtgtctgttcctccgcatcaacagctgcagctctgagcgcgctgcgccccttgggtgcacacccctcga	1552
QY	1509	gagggacagcccaaacacccagagagccctctgtagccgcgcacatcccgcgccaaagacagatgtagacat	1568
Db	1553	gagggacagcccaaacacccagagagccctctgtagccgcgcacatcccgcgccaaagacagatgtagacat	1612
QY	1569	gccaagagacccgcaacaaatccgcgtcgagcgctgtctccctaaaggaaattccctgcatactgacagct	1628
Db	1613	gccaagagacccgcaacaaatccgcgtcgagcgctgtctccctaaaggaaattccctgcatactgacagct	1672
QY	1629	ggctcagcatctcccaagaaagaaatggtgtgtcccccacaccccaagctcagctcgtatgagagag	1688
Db	1673	ggctcagcatctcccaagaaagaaatggtgtgtcccccacaccccaagctcagctcgtatgagagag	1732
QY	1689	tgaagcccaacagaaaccttaacgtctgagagaggtgcagacacgtctaaagtctaaagaaacatacgaagag	1748
Db	1733	tgaagcccaacagaaaccttaacgtctgagagaggtgcagacacgtctaaagtctaaagaaacatacgaagag	1792
QY	1749	gcaaggtgtgcgccccttccctctttaaaagagcccgctgtgtctctggagagaaatccctcagatccc	1808
Db	1793	gcaaggtgtgcgccccttccctctttaaaagagcccgctgtgtctctggagagaaatccctcagatccc	1852
QY	1809	actaaaggtctaaaggtgtctgaaagagagcgacaaagcgaacaaagagataagggcatctggggtctat	1868
Db	1853	actaaaggtctaaaggtgtctgaaagagagcgacaaagcgaacaaagagataagggcatctggggtctat	1912
QY	1869	gccaacatacccaacaggttgcctgtctgcctcccgagcgctctttccatgtgcacatccataatgctcct	1928
Db	1913	gccaacatacccaacaggttgcctgtctgcctcccgagcgctctttccatgtgcacatccataatgctcct	1972
QY	1929	gtctcccaacttcccaactcgtcttcccctccctccctcccgagcgtcgtcttgtgggctcaagcgctg	1988
Db	1973	gtctcccaacttcccaactcgtcttcccctccctccctcccgagcgtcgtcttgtgggctcaagcgctg	2032
QY	1989	taccatccggagaggtctgatactgtagatctgtggagagctccctctaaagagatagagaaacccagga	2048
Db	2033	taccatccggagaggtctgatactgtagatctgtggagagctccctctaaagagatagagaaacccagga	2092
QY	2049	ggcctcgacaccaaagtctcaagaaagctctgtagctgagccatccatccggcacaatactctgtgtc	2108
Db	2093	ggcctcgacaccaaagtctcaagaaagctctgtagctgagccatccatccggcacaatactctgtgtc	2152
QY	2109	tctctcatccacttggaaacacatttattaagacatcgataaataagtgtagcgctcgttggggtatata	2168
Db	2153	tctctcatccacttggaaacacatttattaagacatcgataaataagtgtagcgctcgttggggtatata	2212
QY	2169	cagatctgacatacagataatagatcctcgagctccagagctttaaagctttaaataaaacaaacag	2228
Db	2213	cagatctgacatacagataatagatcctcgagctccagagctttaaagctttaaataaaacaaacag	2272
QY	2229	aaacacaaacaaacttggatctcaaaagagaaataatgataagctgacaaagcagcacaaagaa	2288
Db	2273	aaacacaaacaaacttggatctcaaaagagaaataatgataagctgacaaagcagcacaaagaa	2331
QY	2289	tttccctcgtctgagaaatcctgagcgtctgatactgtagagaaacgggtctaaccaagctbaaagttccc	2348
Db	2332	tttccctcgtctgagaaatcctgagcgtctgatactgtagagaaacgggtctaaccaagctbaaagttccc	2391
QY	2349	gagagacatgctctgtagagagcagaagggcacaacacgacagcgtgtgaatgcgtgtgtgtgat	2408
Db	2392	gagagacatgctctgtagagagcagaagggcacaacacgacagcgtgtgaatgcgtgtgtgtgat	2451
QY	2409	tttgtgtgagtgaagctcgtgttttcgcacacttgatgtgggctggggtttgtctccgtgggctgaa	2468
Db	2452	tttgtgtgagtgaagctcgtgttttcgcacacttgatgtgggctggggtttgtctccgtgggctgaa	2511

|||||
Db 837 tgcagagcccccacgaggaagagagctgcacaaagtgcagagctccggagaaagatctgct 896
QY 799 ctttgaaggtctctctgcagagctgcgggggagagatgcagagctgcggagaaacaaaccc 858
Db 897 ctttgaaggtctctctgcagagctgcgggggagagatgcagagctgcggagaaacaaaccc 956
QY 859 cagccgagagctgcgggagagagagatctctctctgcagagctgcgggagagagagagag 918
Db 957 cagccgagagagctgcgggagagagagatctctctctgcagagctgcgggagagagagag 1016
QY 919 cctaacatgtctcaaaagagatcacaagctctgcacaaagatcattctcaattcagagagctgc 978
Db 1017 cctaacatgtctcaaaagagatcacaagctctgcacaaagatcattctcaattcagagagctgc 1076
QY 979 ccatcagagagagagagatctctctctgcagagagagagagagagagagagagagagag 1038
Db 1077 ccatcagagagagagagatctctctctgcagagagagagagagagagagagagagagag 1136
QY 1039 tcaaacagagctgcacaaag 1098
Db 1137 tcaaacagagctgcacaaag 1196
QY 1099 tggag 1158
Db 1197 tggag 1256
QY 1159 acaatgcag 1218
Db 1257 acaatgcag 1316
QY 1219 tctctcccccag 1278
Db 1317 tctctcccccag 1376
QY 1279 aattgcag 1338
Db 1377 aattgcag 1436
QY 1339 tctctgtctctgaag 1398
Db 1437 tctctgtctctgaag 1496
QY 1399 cccag 1458
Db 1497 cccag 1556
QY 1459 tgttcgag 1518
Db 1557 tgttcgag 1616
QY 1519 gacccag 1578
Db 1617 gacccag 1676
QY 1579 gacaaag 1638
Db 1677 gacaaag 1736
QY 1639 cctcag 1698
Db 1737 cctcag 1796
QY 1699 gactcttaag 1758
Db 1797 gactcttaag 1856
QY 1759 ccttccctttaaaag 1818
Db 1857 ccttccctttaaaag 1916
QY 1819 caaggtgtggaag 1878
|||||

Db 1917 caaggtgtggaag 1976
QY 1879 ccaaggtgtggaag 1938
Db 1977 ccaaggtgtggaag 2036
QY 1939 tcccaaggtgtggaag 1998
Db 2037 tcccaaggtgtggaag 2096
QY 1999 caggtgtggaag 2058
Db 2097 caggtgtggaag 2156
QY 2059 aaggtgtggaag 2118
Db 2157 aaggtgtggaag 2216
QY 2119 ttggaag 2178
Db 2217 ttggaag 2276
QY 2179 tcaagatag 2238
Db 2277 tcaagatag 2336
QY 2239 aattgtgag 2298
Db 2337 aattgtgag 2395
QY 2299 tgaag 2358
Db 2396 tgaag 2455
QY 2359 gtcgtgtggaag 2418
Db 2456 gtcgtgtggaag 2515
QY 2419 taaggtgtggaag 2478
Db 2516 taaggtgtggaag 2575
QY 2479 gtcgtgtggaag 2538
Db 2576 gtcgtgtggaag 2635
QY 2539 gcaactatatttctgtgtacacatctatctcaaaagagagagagagagagagagagag 2598
Db 2636 gcaactatatttctgtgtacacatctatctcaaaagagagagagagagagagagagag 2695
QY 2599 ctgtttatag 2658
Db 2696 ctgtttatag 2755
QY 2659 taaggtgtggaag 2718
Db 2756 taaggtgtggaag 2815
QY 2719 tcaaggtgtggaag 2777
Db 2816 tcaaggtgtggaag 2834
RESULT 6
X78807
ID X78807 standard; CDNA; 3093 BP.
XX
AC X78807;
XX
DT 06-SEP-1999 (first entry)
XX
DE Human nmr7 CDNA;
XX

KM nmK7-1; nuclear trans-acting receptor protein; human; regulator
KM identification; downstream target gene; cell proliferation;
KM cell development; ss.

OS Homo sapiens.

FH	Key	Location/Qualifiers
FT	CDS	275..1676

```
ET /product= "nnr7"
ET /note= "Partial sequence, no start codon given"
```

PN W09931129-A1.

PD 24-JUN-1999

PF 11-DEC-1998; 98WO-US26364.

PR 14-OCT-1998; 98US-0104251.

XX

XX

XX

DR P-PSDB; Y25410.

PT DNA encoding hu

PS Claim 9; Fig 1A-C; 80pp; English.

CC This invention describes the ISO1

This invention describes the isolation of the novel human nuclear receptors nNR7 or nNR7-1. The nNR7 and nNR7-1 proteins are useful in the identification of downstream target genes and ligands regulating its activity. The nuclear receptor is involved in the regulation of *in vivo* cell proliferation and/or cell development. The nNR7 and nNR7-1 polynucleotides, expression vectors and host cells are useful for the recombinant production of the protein.

Sequence 3093 BP; 774 A; 786 C; 843 G; 690 T; 0 other;

Query Match	92.3%	Score 2586.2;	DB 20;	Length 3093;
Best Local Similarity	99.5%	Pred. No. 0;		
Matches 2605; Conservative	0;	Mismatches 13;	Indels 1;	Gaps 1.

OY	139	agttggtcttcacaaacccatccaaagagccccaagaaagaaacactgtgaagtgtagaacccaagaataa	198
Db	334	acttacccacaaagcagctccaaagagccccaagaaacaaacactgtgaagtgtagaacccaagaataa	399
OY	139	gtctggaacccatgctgctgactctttgtacaactgtgtgaagacacagaagctctgtctccctgtgaagaagccca	258
Db	394	gtctggaacccatgctgctgactctttgtacaactgtgtgaagacacagaagctctgtctccctgtgaagaagccca	455
OY	259	gtgtccaaagccagatgagagaaagctcggaaggtgccccaatctgcgctgtatgtgtgtgtgtgaagaag	318
Db	454	gtgtccaaagccagatgagagaaagctcggaaggtgccccaatctgcgctgtatgtgtgtgtgtgaagaag	513
OY	319	ccaactgctatcaactccaatgtcatgtacatgtgtgaagagatgtgaagagctcttttcacagagag	378
Db	514	ccaactgctatcaactccaatgtcatgtacatgtgtgaagagatgtgaagagctcttttcacagagag	573
OY	379	ccatgtgaaacccgaaacgccccggctgtgaggtgtgcgcccttcgcgaaagagggcgccctgtgagaaagtcaacc	438
Db	574	ccatgtgaaacccgaaacgccccggctgtgaggtgtgcgcccttcgcgaaagagggcgccctgtgagaaagtcaacc	633
OY	439	ggaagaaaccccgccgacagatgtgcacaaagcctgtgcgcctgtgcgacaaagtgtgcctgtgaagacggtcatga	498
Db	634	ggaagaaaccccgccgacagatgtgcacaaagcctgtgcgcctgtgcgacaaagtgtgcctgtgaagacggtcatga	693
OY	499	aggaagagagatgatatgtgccgaaacgaagccgctgtgagagagagagcggtgccttatacaagcggaa	558

Dh	694	agaaagagaaatgatactacatctccgaagaaagacggttggaagaaagacgagcccttgatacgaagcga	753
Qy	559	agaaaaagatgaaacacggaaacagagactacaaagacatctggagatgacaggggctgtgaacaagaagaaagac	618
Dh	754	agaaaaagatgaaacacggaaacagagactacaaagacatctggagatgacaggggctgtgaacaagaagaaagac	813
Qy	619	ggaatgatactaaagaaagacatgataagctacaatgaaacacttgacaatacctctcc	678
Dh	814	ggaatgatactaaagaaagacatgataagctacaatgaaacacttgacaatacctctcc	873
Qy	679	attcaaaataattcccgctccacaaagggttgcttaagcatgtgcctgcagattggccaagatctc	738
Dh	874	attcaaaataattcccgctccacaaagggttgcttaagcatgtgcctgcagattggccaagatctc	933
Qy	739	tgcagagcccccacgcgaagaaagaaagctgcacaaagtgaagccaagatcccgaaaaagatactgtgct	798
Dh	934	tgcagagcccccacgcgaagaaagaaagctgcacaaagtgaagccaagatcccgaaagatctgtgct	993
Qy	799	ctttgaaaggtctctctctgcagctcgcgggggagaaatgacagatgctctgaaactacaacccc	858
Dh	994	ctttgaaaggtctctctctgcagctcgcgggggagaaatgacagatgctctgaaactacaacccc	1053
Qy	859	cagccggaagaaagtgaggaaagaaagatactctccctgtgcacaaatgctgacatgataa	918
Dh	1054	cagccggaagaaagtgaggaaagaaagatactctccctgtgcacaaatgctgacatgataa	1113
Qy	919	ccatacatgtctcaaaagagcatatcatagctcttgccaaaatcatctccatactcaagagaaatgctgc	978
Dh	1114	ccatacatgtctcaaaagagcatatcatagctcttgccaaaatcatctccatactcaagagaaatgctgc	1173
Qy	979	ccatccagagaaacagatactccctctgtgaagggggcgcttccgaagctgtgtcaaatgaaat	1038
Dh	1174	ccatccagagaaacagatactccctctgtgaagggggcgcttccgaagctgtgtcaaatgaaat	1233
Qy	1039	tcaacaacaagtcttcaaaagcgggaagaaacgggaacccgggaaggtgcgcgggtgtccatactct	1098
Dh	1234	tcaacaacaagtcttcaaaagcgggaagaaacgggaacccgggaaggtgcgcgggtgtccatactctct	1293
Qy	1099	tggaaagaaacacgcgaagctgcgtcttcacaaagaaactctaaatgaaagccaatgtctgaataatcaact	1158
Dh	1294	tggaaagaaacacgcgaagctgcgtcttcacaaagaaactctaaatgaaagccaatgtctgaataatcaact	1353
Qy	1159	acatgctgaaagaaagctgcagctgcatagtaaagaaagaaatgtgctgatacagagccaatctccc	1218
Dh	1354	acatgctgaaagaaagctgcagctgcatagtaaagaaagaaatgtgctgatacagagccaatctccc	1413
Qy	1219	tctctctcccaagaaacccgcgaaggtgtcctctgcacaaacccgggtgtgtgaacaagatctgcagaagac	1278
Dh	1414	tctctctcccaagaaacccgcgaaggtgtcctctgcacaaacccgggtgtgtgtgaacaagatctgcagaagac	1473
Qy	1279	aattccgcaatatactctgaaaatctccataatgaaatgtcaatctgggcccacagactctgataaagt	1338
Dh	1474	aattccgcaatatactctgaaaatctccataatgaaatgtcaatctgggcccacagactctgataaagt	1533
Qy	1339	tctgtgtctctgaaagatcatatgtgctataccgaagactccgcagacataatgctcagaaca	1398
Dh	1534	tctgtgtctctgaaagatcatatgtgctataccgaagactccgcagacataatgctcagaaca	1593
Qy	1399	ccccagcgctgcgtgcgcgatctccaaagaaatacaaacaccttggctaaagccctctatgaagaagt	1458
Dh	1594	ccccagcgctgcgtgcgcgatctccaaagaaatacaaacaccttggctaaagccctctatgaagaagt	1653
Qy	1459	agttctgcgcacatacaaggttaagcttgagcgggtgcgcctctgggttgaaacctccgaaagagcagcca	1518
Dh	1654	agttctgcgcacatacaaggttaagcttgagcgggtgcgcctctgggttgaaacctccgaaagagcagcca	1713
Qy	1519	gaacccagagacatctcgaagccgcgaacatcccgggccaaagacaaatgaaacatgctccaaagacc	1578
Dh	1714	gaacccagagacatctcgaagccgcgaacatcccgggccaaagacaaatgaaacatgctccaaagacc	1773
Qy	1579	gacaatagctctgtgctgcgtctctcctaaagaaatctcgtcatgacagctgtgctaaacatt	1638
Dh	1774	gacaatagctctgtgctgcgtctctcctaaagaaatctcgtcatgacagctgtgctaaacatt	1833

PA (SALK) SALK INST BIOLOGICAL STUDIES.

XX Blumberg B, Evans RM;

DR WPI: 1999-419349/35.

DR P-PSDB: Y21799.

XX New steroid and xenobiotic receptor, used to identify modulators for
PT controlling metabolism of steroids and xenobiotics, e.g. reducing
PT their toxicity

PS Claim 10: Fig 1A: 83pp; English.

CC The invention relates to a novel nuclear receptor polypeptide, designated
CC SXR (steroid and xenobiotic receptor). SXR (1) forms a heterodimer with
CC retinoid X receptor (RXR), (11) binds to a direct or inverted repeat
CC response element motif based on the half-site AGGTCA, (111) activates
CC transcription through response elements present in steroid-inducible P450
CC genes, in response to a wide variety of natural and synthetic steroid
CC hormones and (iv) is prominently expressed in liver and intestine. SXR
CC regulates expression of catabolic enzymes, in response to many different
CC steroids, and thus affects metabolism. SXR is a broad specificity, low-
CC affinity receptor for reducing excessive levels of steroids in the
CC circulation. (Antagonists of SXR are used to regulate metabolism of
CC steroids particularly phytoestrogens or calcium-channel blockers, to
CC reduce steroid toxicity in subjects being treated with steroids, e.g. in
CC cases of tuberculosis (treated with rifampin and related compounds),
CC breast cancer (treated with tamoxifen, raloxifen etc.) or osteoporosis
CC (treated with Vitamin K), or to slow metabolism of therapeutic steroids.
CC Also, modulating endogenous SXR is used to treat disease, particularly
CC an agonist is used where endogenous steroid levels are excessive (e.g.
CC Cushing syndrome; virilism and hirsutism in women; polycystic ovarian
CC disease; 11 beta-, 17- or 21-hydroxylase deficiency; 3 beta-hydroxysteroid
CC dehydrogenase deficiency, or breast, colorectal or prostatic cancer),
CC while antagonists are used where endogenous steroid levels are too low.
CC Cells that express SXR are used to identify compounds likely to be
CC involved in undesirable drug interactions. Antibodies specific for SXR
CC are used in immunohistochemical testing for studying distribution/
CC expression density of SXR, also for diagnosis and therapeutically as
CC antagonist. The present sequence represents the longest SXR cDNA clone
CC encoding the SXR polypeptide.

CC Sequence 2068 BP; 520 A; 541 C; 586 G; 420 T; 1 other;

Query Match 52.3%; Score 1466.6; DB 20; Length 2068;

Best Local Similarity 98.8%; Pred. No. 5.1e-294;

Matches 1508; Conservative 0; Mismatches 15; Indels 3; Gaps 3;

QY 139 agtggctcaaaacccatccagaagcccaagaacaaacctggagtgtagaaccacaaga 198
DB 545 actaacacacaaagagccaaagagcccaagaacaaacctggagtgtagaaccacaaga 604
QY 199 gctggaacacatgtgacttctgtaacctgtgagagacacagatctgttcttgaaagcca 258
DB 605 gctggaacacatgtgacttctgtaacctgtgagagacacagatctgttcttgaaagcca 664
QY 259 gcttcaacagcagatgagagagtcggaggtcccaaatctgcgtgtaatgaggaagag 318
DB 665 gcttcaacagcagatgagagagtcggaggtcccaaatctgcgtgtaatgaggaagag 724
QY 319 ccaactgctcactcaatcagatgacatgtgaagagatcagaaggtcttttaagagag 378
DB 725 ccaactgctcactcaatcagatgacatgtgaagagatcagaaggtcttttaagagag 784
QY 379 ccatgaacacgaacgcgcgcgtgagtgcccttcggaaaggcgccctgagatcaccc 438
DB 785 ccatgaacacgaacgcgcgcgtgagtgcccttcggaaaggcgccctgagatcaccc 844
QY 439 ggaagacccggcgacagtgccagagcctgcgcctgagcaagtgccggagagagcgatga 498
DB 845 ggaagacccggcgacagtgccagagcctgcgcctgagcaagtgccggagagagcgatga 904

QY 499 agaaagagatgatcatgtccagagagcggtgagagagagcggtctgtatcaagcgga 558
DB 905 agaaagagatgatcatgtccagagagcggtgagagagagcggtctgtatcaagcgga 964
QY 559 agaaagatgaaag 618
DB 965 agaaagatgaaag 1024
QY 619 ggaatgagatcagaag 678
DB 1025 ggaatgagatcagaag 1084
QY 679 attcaagaattcccggtctgcaggggtgcttagcagtggtgagagagagagagagag 738
DB 1085 attcaagaattcccggtctgcaggggtgcttagcagtggtgagagagagagagagag 1144
QY 739 tgcagagcccatcag 798
DB 1145 tgcagagcccatcag 1204
QY 799 ctgtgaaggtctctctgc-agctcgaggaggagagagagagagagagagagagagag 857
DB 1205 ctgtgaaggtctctctgcag 1264
QY 858 ccaagcagacagtggtcgag 917
DB 1265 ccaagcagacagtggtcgag 1323
QY 918 acctaatgttcaagagatcatcatcagagagagagagagagagagagagagagagag 977
DB 1324 acctaatgttcaagagatcatcatcagagagagagagagagagagagagagagagag 1383
QY 978 cccatgag 1037
DB 1384 cccatgag 1443
QY 1038 ttcaacacag 1097
DB 1444 ttcaacacag 1503
QY 1098 ttggaagacag 1157
DB 1504 ttggaagacag 1563
QY 1158 tacatgctgaaag 1217
DB 1564 tacatgctgaaag 1623
QY 1218 ctcttctcccaagacagcag 1277
DB 1624 ctcttctcccaagacagcag 1683
QY 1278 caattcgcaattactctgaag 1337
DB 1684 caattcgcaattactctgaag 1743
QY 1338 ttcttggctcgtgaag 1397
DB 1744 ttcttggctcgtgaag 1803
QY 1398 acccagcggtctgctgagatcagaagacatacaccccttgcagagagagagagagag 1457
DB 1804 acccagcggtctgctgagatcagaagacatacaccccttgcagagagagagagagag 1863
QY 1458 ttgttcgagacacacag 1517
DB 1864 ttgttcgagacacacag 1922
QY 1518 agaacag 1577
DB 1923 agaacag 1982
QY 1578 ggaacag 1637

ID X59966 standard; DNA: 1305 BP.
 AC X59966;
 XX 04-AUG-1999 (first entry)
 DT 04-AUG-1999 (first entry)
 DE DNA encoding an intranuclear receptor protein.
 XX Human; intranuclear receptor protein; drug development; diagnosis;
 KW treatment; ss.
 XX Homo sapiens.
 OS JP1127872-A.
 PN 18-MAY-1999.
 PD 07-AUG-1998; 98JP-0224172.
 PF 11-AUG-1997; 97JP-0230335.
 PR (NIBS) JAPAN TOBACCO INC.
 PA WPI: 1999-350330/30.
 DR P-PSDB; Y15931.
 XX New intranuclear receptor protein - useful for drug development and
 PT diagnosis and treatment of disease
 PS Claim 4; Page 22-23; 38pp; Japanese.
 CC The present sequence encodes a human intranuclear receptor protein.
 CC The nucleic acid sequence was isolated from a human adult cDNA
 CC library using a swedish ANO23 derived probe. The protein can
 CC be used for the development of drugs and diagnosis and treatment
 CC of various diseases.
 SQ Sequence 1305 BP; 304 A; 356 C; 386 G; 259 T; 0 other;
 Query Match 46.6%; Score 1305; DB 20; Length 1305;
 Best Local Similarity 100.0%; Pred. No. 1.1e-260;
 Matches 1305; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 421 caggggctgacagagagcagcgatgatgatcagaggagctgacgctcagatgaa 480
 Qy 657 accttgaacatcctcttccattccaagaattccggctgcaggggtgcttaagcag 716
 Db 481 accttgaacatcctcttccattccaagaattccggctgcaggggtgcttaagcag 540
 Qy 717 ggcctgagcttgcagagatctctgagggccctgaggggaagaagctccaaagtgaac 776
 Db 541 ggcctgagcttgcagagatctctgagggccctgaggggaagaagctccaaagtgaac 600
 Qy 777 caggctcggagaagatctgctcttgaaggtctctctgagctgcggggggaagatggc 836
 Db 601 caggctcggagaagatctgctcttgaaggtctctctgagctgcggggggaagatggc 660
 Qy 837 agtctcggagaacacaaaccccgagcgagatggcggggaagaagatctctcctgctg 896
 Db 661 agtctcggagaacacaaaccccgagcgagatggcggggaagaagatctctcctgctg 720
 Qy 897 ccccaatgctgacatgctacacatgcttcaagaagcagcagcttgcaggaagtc 956
 Db 721 ccccaatgctgacatgctacacatgcttcaagaagcagcagcttgcaggaagtc 780
 Qy 957 atctctactcagggagctgcccacgcagagacagatccctgctgtaaggggccgct 1016
 Db 781 atctctactcagggagctgcccacgcagagacagatccctgctgtaaggggccgct 840
 Qy 1017 ttgagagctggttaactgagatccaacacagtgcttaagcggagagctggaacccgggag 1076
 Db 841 ttgagagctggttaactgagatccaacacagtgcttaagcggagagctggaacccgggag 900
 Qy 1077 ttgagagctggttaactgagatccaacacagtgcttaagcggagagctggaacccgggag 1136
 Db 901 ttgagagctggttaactgagatccaacacagtgcttaagcggagagctggaacccgggag 960
 Qy 1137 gaggccatgctgaatccaactacatgctgaagaagctgacgtcatgagagagatg 1196
 Db 961 gaggccatgctgaatccaactacatgctgaagaagctgacgtcatgagagagatg 1020
 Qy 1197 gtgctgatgagagcctctccctcttccccaagaccccgagtggtgtagaacacgc 1256
 Db 1021 gtgctgatgagagcctctccctcttccccaagaccccgagtggtgtagaacacgc 1080
 Qy 1257 gtgctgatgagagcctctccctcttccccaagaccccgagtggtgtagaacacgc 1316
 Db 1081 gtgctgatgagagcctctccctcttccccaagaccccgagtggtgtagaacacgc 1140
 Qy 1317 cggcccaagctgtctataggttctgtctcctgaagatcatgctatgctacacgagctc 1376
 Db 1141 cggcccaagctgtctataggttctgtctcctgaagatcatgctatgctacacgagctc 1200
 Qy 1377 cggagatcatgctatgagacaccccgagcgctgctgagatcatcagagatcatcaccctt 1436
 Db 1201 cggagatcatgctatgagacaccccgagcgctgctgagatcatcagagatcatcaccctt 1260
 Qy 1437 gctacgcccctcatgagagatggttcgcatcacaggtgagctga 1481
 Db 1261 gctacgcccctcatgagagatggttcgcatcacaggtgagctga 1305
 RESULT 12
 X59972
 ID X59972 standard; DNA: 644 BP.
 AC X59972;
 XX 04-AUG-1999 (first entry)
 DT 04-AUG-1999 (first entry)
 DE SEQ ID 12 of JP1127872.
 XX Human; intranuclear receptor protein; drug development; diagnosis;
 KW treatment; ss.
 XX

OS Homo sapiens.
 XX JP1127872-A.
 XX 18-MAY-1999.
 XX 07-AUG-1998; 98JP-0224172.
 XX 11-AUG-1997; 97JP-0230335.
 XX (NINS) JAPAN TOBACCO INC.
 XX WPI: 1999-350330/30.
 DR New intranuclear receptor protein - useful for drug development and
 PT diagnosis and treatment of disease
 XX Disclosure: Page 32, 38pp; Japanese.
 XX The specification describes a human intranuclear receptor protein.
 CC The nucleic acid sequence was isolated from a human adult CDNA
 CC library using a swellfish ANO23 derived probe. The protein can
 CC be used for the development of drugs and diagnosis and treatment
 CC of various diseases. The present sequence was used in the course of
 CC the invention.
 XX Sequence 644 BP; 175 A; 161 C; 193 G; 115 T; 0 other;

Query Match 15.2%; Score 425.2; DB 20; Length 644;
 Best Local Similarity 98.2%; Pred. No. 4.5e-79;
 Matches 430; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 139 agtggcttcaaacatcccaagagcccaagagcaaacctgaggtgagcccaagaa 198
 DB 138 acttaccacacagagcagcagagagcccaagagcaaacctgaggtgagcccaagaa 257
 QY 199 gctggaacacatgctgactgtgacactgtgagagacacagagctgtctcctggaagccca 258
 DB 258 gctggaacacatgctgactgtgacactgtgagagacacagagctgtctcctggaagccca 317
 QY 259 gtgtcaacgcagatgagagagtcgagagtcgcccaaatctgcctgtatgtgagagcaagg 318
 DB 318 gtgtcaacgcagatgagagagtcgagagtcgcccaaatctgcctgtatgtgagagcaagg 377
 QY 319 ccactggcctactactcaatgtcatgacatgtgaaagatgcaaggcttttcaggaagg 378
 DB 378 ccactggcctactactcaatgtcatgacatgtgaaagatgcaaggcttttcaggaagg 437
 QY 379 ccattgaacacagcagcccgctgaggtgagcccttcggaagagcgccctgcaagatcaccc 438
 DB 438 ccattgaacacagcagcccgctgaggtgagcccttcggaagagcgccctgcaagatcaccc 497
 QY 439 ggaagaccccgagcagatgccaagcctgcgcctgcgcaagatgctgagagcggaatga 498
 DB 498 ggaagaccccgagcagatgccaagcctgcgcctgcgcaagatgctgagagcggaatga 557
 QY 499 aagaagagatgctatgttcgagagcgccgtgagagagagcgagcgcttgatcaagcgga 558
 DB 558 aagaagagatgctatgttcgagagcgccgtgagagagagcgagcgcttgatcaagcgga 617
 QY 559 aagaagatgcaagcagag 576
 DB 618 aagaagatgcaagcagag 635

RESULT 13
 V03129
 ID V03129 standard; CDNA; 1960 BP.
 AC V03129;
 DX 11-JUN-1998 (first entry)

XX CDNA encoding rat vitamin D receptor (VDR).
 DE Rat; vitamin D receptor; isoform protein; VDR1; VDR0; diagnosis;
 XX dominant negative receptor; signal transmission channel;
 KW bone density disorder; screening; ds.
 OS Rattus rattus.
 XX Key Location/Qualifiers
 FT CDS .12..1283
 FT /tag- a
 FT /transl_except- (pos:1089..1091, aa:gly)
 XX MO9747172-A1.
 XX 18-DEC-1997.
 XX 10-JUN-1997; 97MO-IB00947.
 XX 10-JUN-1996; 96JP-0194179.
 XX (CHUS) CHUGAI PHARM CO LTD.
 XX (CHUS) CHUGAI SEIYAKU KK.
 XX Kato S, Ueno K;
 XX WPI: 1998-051917/05.
 DR P-PSDB; M47509.
 XX DNA encoding a vitamin D receptor isoform protein - useful for bone
 PT density determination and for screening substances for vitamin D
 PT activity
 XX Disclosure: Fig 1; 46pp; Japanese.
 XX A novel CDNA sequence encodes the rat vitamin D receptor isoform
 CC protein (VDR1). The isoform differs from the normal receptor
 CC (VDR0), which is encoded by the present sequence, in having the
 CC vitamin D response element curtailed by 86 residues, and having an
 CC extra 19 residues inserted at the C-terminal of this element. It
 CC acts as a dominant negative receptor in the vitamin D signal
 CC transmission channel.
 CC The isoform protein can be used to diagnose bone density disorders,
 CC and screen for substances having potential vitamin D-like activity.
 XX Sequence 1960 BP; 432 A; 620 C; 497 G; 411 T; 0 other;

Query Match 10.0%; Score 278.8; DB 19; Length 1960;
 Best Local Similarity 55.6%; Pred. No. 1.1e-48;
 Matches 629; Conservative 0; Mismatches 482; Indels 21; Gaps 4;

QY 288 ccccaaatctgcctgtatgtgtgagagcaagccactggtatcaattcaatgtcatgaca 347
 DB 72 ccccgagctcgtgtgaggtgtgagagccgagccagcagcttcaattcaatgtcatgaca 131
 QY 348 ttgtgaagatgcaagagcttttcagagagggccatgaagcacaagcccgctgagagtc 407
 DB 132 ttgtgaagatgcaagagcttttcagagagggccatgaagcacaagcccgctgagagtc 191
 QY 408 cccctcggagagcgccgctgcgagatcacccggagagacccggcagagtgccagcgctgc 467
 DB 192 cccctc---aatgagagatgcccatcaccaagagacacccggcagagtgccagcgctgc 248
 QY 468 cgcctgcgcaagtgctcggagagcgccatgaagagagagatgcatgttcgagcgagcc 527
 DB 249 cgcctgcgcaagtgctcggagagcgccatgaagagagagatgcatgttcgagcgagcc 308
 QY 528 gtgagagagagcgccgcttgatcagcagcgaaagaaagtgaaagcgagagatccgca 587
 DB 309 gtacagcgttagagagagagatgaatgaagaaagagagagagcgcttgagagagcagt 368

[illegible]

```
Search completed: February 18, 2001, 16:55:50
Job time: 8630 sec
```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 18, 2001, 16:45:23 ; Search time 6591.86 Seconds
(without alignments)
2175.401 Million cell updates/sec

Title: US-09-143-828-3

Perfect score: 2802
Sequence: 1 tgaatcgtggcctcgtcgtg.....aaaaaaaaaaaaaaaaaaaaa 2802

Scoring table: IDENTITY_NDC
Gapop 10.0 , Gapext 1.0

Searched: 118133 segs, 2558875100 residues

Total number of hits satisfying chosen parameters: 2236266

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenDb1: *
1: gb_da1: *
2: gb_da2: *
3: gb_da3: *
4: gb_da4: *
5: gb_da5: *
6: gb_da6: *
7: gb_da7: *
8: gb_da8: *
9: gb_da9: *
10: gb_da10: *
11: gb_da11: *
12: gb_da12: *
13: gb_da13: *
14: gb_da14: *
15: gb_da15: *
16: gb_da16: *
17: gb_da17: *
18: gb_da18: *
19: gb_da19: *
20: gb_da20: *
21: gb_da21: *
22: gb_da22: *
23: gb_da23: *
24: gb_da24: *
25: gb_da25: *
26: gb_da26: *
27: gb_da27: *
28: gb_da28: *
29: gb_da29: *
30: gb_da30: *
31: gb_da31: *
32: gb_da32: *
33: gb_da33: *
34: gb_da34: *
35: gb_da35: *
36: gb_da36: *
37: gb_da37: *
38: gb_da38: *
39: gb_da39: *
40: gb_da40: *
41: gb_da41: *
42: gb_da42: *
43: gb_da43: *

44: em_htg8: *
45: em_htg9: *
46: em_htg10: *
47: em_htg11: *
48: em_htg12: *
49: em_htg13: *
50: em_htg14: *
51: em_htg15: *
52: em_htg16: *
53: em_htg17: *
54: em_htg18: *
55: em_htg19: *
56: em_htg20: *
57: em_htg21: *
58: em_htg22: *
59: em_htg23: *
60: em_htg24: *
61: em_htg25: *
62: em_htg26: *
63: em_htg27: *
64: em_htg28: *
65: em_htg29: *
66: em_htg30: *
67: em_htg31: *
68: em_htg32: *
69: em_htg33: *
70: em_htg34: *
71: em_htg35: *
72: em_htg36: *
73: em_htg37: *
74: em_htg38: *
75: em_htg39: *
76: em_htg40: *
77: em_htg41: *
78: em_htg42: *
79: em_htg43: *
80: em_htg44: *
81: em_htg45: *
82: em_htg46: *
83: em_htg47: *
84: em_htg48: *
85: em_htg49: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2802	100.0	2802	10 AF084644	AF084644 Homo sapi
2	2651.2	94.6	2905	10 AF084645	AF084645 Homo sapi
3	2591.2	92.5	4448	52 HSAJ9936	AJ009936 Homo sapi
4	2359.2	84.2	4337	52 HSAJ9937	AJ009937 Homo sapi
5	1866.6	66.6	2146	10 AF061056	AF061056 Homo sapi
6	1398.8	49.9	160185	69 AC069444	AC069444 Homo sapi
7	1046	37.3	1895	3 AF182217	AF182217 Oryctolagus
8	991.4	35.4	1601	3 AF188476	AF188476 Oryctolagus
9	974.2	34.8	1709	11 AF031814	AF031814 Mus muscu
10	967.8	34.5	1755	11 AF151377	AF151377 Rattus no
11	335	12.0	365	77 G36928	G36928 SHGC-56597
12	284	10.1	1340	4 AF276753	AF276753 Gallus ga
13	280.8	10.0	1377	11 MGSVDR	D31969 Mouse mRNA
14	278.8	10.0	2043	11 BATPDHRO3	J04147 Rat 1.25-cl
15	275	9.8	2494	4 AF011356	AF011356 Gallus ga
16	272.8	9.7	1284	10 AF026260	AF026260 Homo sapi
17	272.8	9.7	1335	53 HSDVDR	X67482 H. sapiens m
18	272.8	9.7	4604	85 HSDVDR	J03258 Human vitam
19	267.8	9.6	1724	4 C7012641	U13641 Cocurnix ja
20	258.6	9.2	1638	4 XLRNAONST	X75163 X. laevis MR
21	239.8	8.6	2429	4 AB037674	AB037674 Paratuberc

[illegible]

1
1
1

Db	241	CTTCTCTCTGGAGAGGCCCATGTCTCAACGAGTGGAGAGTCGGAGGTCCCAAACTGTGC	300
Qy	301	gtgtatgtggacaagcccaactgctcatcacttcaatgtcatgtacatgtgaagatgca	360
Db	301	GTGTATGTGGACAAGCCCACTGCTCATCACTTCAATGTCAATGACATGTGAAGATGCA	360

421 ggcgcctgcagatccaccccggaagacccgcgcgaatgfcagagcctgcgcctgcgaagt 480

QY 481 gccctgagagcgcgcatgaagaagagctgcatcattccgacgagcgctggagagagc 540
Db 481 gccctgagagcgcgcatgaagaagagctgcatcattccgacgagcgctggagagagc 540
QY 541 gggccttgatcagcggaagaagtgaaacggaacaggaactcagccactggagtgcag 600

QY	601	gagctacagagtagagcagcgatgtagtgcaggagctgtagtcagatcatgaataacct	660
Db	601	ggcttgcacagagacagcggcgatgatgatctcagggaactgatggacgctcagatgaaataacct	660
QY	661	ttgacactacactcttcccatcttcaagaattctccgcctgcacagaggtgctcttaacatgagct	720

Db	661	TTGACACATACCTTCTCCCATTTCAAGATTTCCGCGCTGCCAGGGGTGCTTACCATGGACT	720
Oy	721	gcgaattgcgaagatctctgcaggcccatcgagggaagaagctgcacatgtgagccagg	780
Db	721	ggcagttgcgagagctcttcgcaggcccatcgagggaagaagctgcacatgtgagccagg	780

Accession	Sequence	Position
Db	TCGGGAAGAAGCTGCTCTTGAAGCTCTCTGCAGCTGGGGGGGAGATGCAATG	840
QY	ttctggaactaaacccccagcgacagtgcggggaagatcttctccctgctgcgcc	900
b	TTGGAACTACAAACCCGACAGCAGTGGGGGAAAGATCTTCTCCCTGCTGCCCC	900

QY 901 acatggtgacatgtcaacctacatgtttcaagcatcatcagcttgcgaagtcalt 960
|||||

Oy	1819	caaggtgtggaagaggccaagcgacacaagaataaggccattctggggtctcatgcccacatc	1878
Db	1922	CAAGGTGTGGAAGAGGACCAAGCACCAAGGATAGGCCATCTGGGCTCATGCCACATAC	1981
Oy	1879	ccaaggtctgtccctccctcagctcttcattctgtaaccttaatagtccgtctcccaact	1938
Db	1982	CCAGGTTGTGTCGTTCCCTCCAGTCTTTCATTGTCATCCCTCAATAGTCTCTGTCCCACT	2041
Oy	1939	tcccactcgcttcccctccctctctccgaagctgtttgtgggtctaaagcctgtatccatcgg	1998
Db	2042	TCCCACACTGCTTCCCTCCCTCTCTTCCGAGCGCTTGTGGGCTCAAGCCGTGATCATCGG	2101
Oy	1999	cagggtgaatgaatcatctgtgtggagatcccttaagagataatgaagaacagagagcctgcac	2058
Db	2102	CAGGTGATAGATATCTGTGGAGTCCCTTAGAGATATGAAAGCCAGAGGCTGTGACC	2161
Oy	2059	aaatgtcagaagaattgtgcatacctcattccggccacatcattctgtctctgcattcca	2118
Db	2162	AAATGTGAGAAGCTTGCGATGAGCCTCATTTCCGGCCACATCATTTCTGTCTCTGCATCCA	2221
Oy	2119	tttgaacaactttatgaagcctgaataatgtatagcctgtgtgtgggtataagcattgac	2178
Db	2222	TTTTGAACACTTTATTAAAGCCTATATAGTAGCTCTGTGGGGTATACAGCATTTGAC	2281
Oy	2179	tcagataatagatccctgtgcctcacagaagttatagttaaaaaaacacagaacaacaac	2238
Db	2282	TCAATATTAATCTCTAGTCCACAGAGTTATAGTTAAABAAAACAACAGAAACAACAAC	2341
Oy	2239	aatttggatcnaaaggaagaaatgatalaagtgcacaaagaacagacaaagaattccctgtg	2298
Db	2342	AATTTGGATCAAAAAGGAGAAATGATATAGTAGCAAAAGCAGACAAAGAAATTTCCCTGTG	2401
Oy	2289	tgggtgtcgtggcgtgtgtgtgcgaagcactgtgtaccagaatgaaagttcccgaggcatga	2358
Db	2402	TGGATGCTGAGCTGTGATGTGGCCAGGCACTGGGTATCCCAATGTAAAGTTCCCGAGGACATGA	2461
Oy	2359	gtctgtatggagcaagaggacaaacacacacagctgtgagttgtgtgtatattgtgtatg	2418
Db	2462	GTCGTATGAGNAGCAAGGGCACAAATCGAGCTGTGATGGCTGTGTGTATTTGGTGTAGG	2521
Oy	2419	taagttcgttttgcacattgtatggggcctvgggttgtctcctggggctgaaatgcvtggtat	2478
Db	2522	TAGGTCTGTTTGGCCACTTGATGTGGGCTGTGGGCTTGTCTCTGGGGCTGGAAATCTGGGAT	2581
Oy	2479	gctctgtgaacaaagcctacgctgtgaacatcagttaaacacacggagagaaagacattatcat	2538
Db	2582	GCTCTGTGACAAGGCTACGCTGACATCACTTAAACACACCGAGAGAAACATTTAAT	2641
Oy	2539	gcacctataatttctgtgtacaacactatctcacaagctaaaggytatgaaggtcctgc	2598
Db	2642	GCACCTTATATTCTGTGTACATCTATCTCAAAAGCTAAAGGGATGAAGTGCCTGC	2701
Oy	2599	cttgtttatagcactcgtgtgataaaaaatttttttgcatttccacaataataactttca	2658
Db	2702	CTTGTTTATTAACCACTGTTGTGATTAATAATTTTTCATTTTCCAAATTAATACCTTTATA	2761
Oy	2659	taaggcatctcacaccccaagaactagtttgggaaaaatgagccctgggttaattgtcaaa	2718
Db	2762	TAAAGCATTCACACCTTAAGAACTATAGTTTGGGAAATGTAGCCCTGGTTTAATGTCAA	2821
Oy	2719	tcaagycanaaagaaataaataatgtatcttttggctaaaaaanaaaaaaanaaaaaa	2778
Db	2822	TCAAGCAAAAGGAATTAATAATGATCTTTGGCTAAAAAANAANAANAANAANAANA	2881
Oy	2779	aaaaaaaaaaaaaaaaaaaaaaaaaa 2802	
Db	2882	AAAAAAAAAAAAAAAAAAAAAAAAAA 2905	

RESULT	3				
HSAJ9936					
LOCUS	HSAJ9936	4448 bp	mRNA	PRI	08-SEP-1999

DEFINITION	Homo sapiens mRNA for nuclear hormone receptor PRRL.
ACCESSION	AJ0009936
VERSION	AJ0009936.1 GI:5852062
KEYWORDS	nuclear hormone receptor; orphan nuclear receptor; PRRL gene.
SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS	Heard,D.J., Holloway,J., Hansen,C., Tommerup,N., Aagaard,L. and Vissing,H.
TITLE	Identification of a novel protein isoform of the human nuclear hormone receptor PXR/SXR and localization to chromosome 3q12.1 - 13.3
JOURNAL	Eur. J. Hum. Genet. In press
REFERENCE	2 (bases 1 to 4448)
AUTHORS	Heard,D.J.
TITLE	Direct Submission
JOURNAL	Submitted (31-JUL-1998) Heard D.J., Novo Nordisk A/S, Molecular Genetics, Novo Alle, DK-2880, Bagsvaerd, DENMARK
FEATURES	Location/Qualifiers
SOURCE	1..4448
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	/chromosome="3"
	/map="q12-q13.3"
	/tissue_type="liver"
	/tissue_type="intestine"
	/tissue_type="colon"
	1771..3144
gene	/gene="PRRL"
	1771..3144
CDS	/gene="PRRL"
	/note="isoform with longer A/B domain, translation initiates at GUG-1771 5 fold less efficiently than at CUG-1840"
	/codon_start=1
	/transl_except=(pos:1771..1773,aa:Met)
	/product="nuclear hormone receptor PRRL-C"
	/protein_id="CAB55490.1"
	/db_xref="GI:5852064"
	/translation="MDPRGEGVAKNTPNSRGPEANLVEPKESNMHTDFVHCEDTE SVPRPSYNADEEVGGPOICVYCDKATGTHFNVTGECCKGFFRRAMKRNARRCPF RKGCETRTKTRROCCARLCLSESKNMNIMSDAVERRALIRKRSERTGTOP LGVGLTEEDRMIRRELMDAQMKTFTTFSHFNRRLPGVLSGCELPESLQASREE AAKSAVRKDCSLKVSILQIGEDGSYWNKPPADSGGKEIFSLPHMAIMSTYMGFE LISPAKVISYPRDPIEDQILKGAPELQOLFNTVFNAETGTCGRSLYLEDTO VAGGQDILPEPMKFHYMKIKRLOHBEVYVMOAISLESPRPGVLQHRVVDLOEOE AATLKSTIECNRPQPAHREFLKIMAMTELRSINAOHTQRLRIDIHPPATLMGE LFGITGS
	1840..3144
CDS	/gene="PRRL"
	/function="transcription factor"
	/note="full length receptor"
	/codon_start=1
	/product="nuclear hormone receptor PRRL-A"
	/protein_id="CAB55489.1"
	/db_xref="GI:5852063"
	/translation="MEVAPKESNMHTDFVHCEDTESYVGRPSYNADEEVGGPOICVYCDK ATGATGAFHNTVTCGCKGFFRRAMKRNARLRCPFRGACITRTTRROCCARLCLSGE KSNMIMSDAVERRALIRKRSERTGTOPLGVLTEEDRMIRRELMDAQMKTFTTFSHFNR RLPGVLSGCELPESLQASREEAAKSAVRKDCSLKVSILQIGEDGSYWNKPPADSGGKEIF SLPHMAIMSTYMGFE LKGAPELQOLFNTVFNAETGTCGRSLYLEDTOVAGGQDILPEPM KFHYMKIKRLOHBEVYVMOAISLESPRPGVLQHRVVDLOEOEPATLKVSICNRPQPAH REFLKIMAMTELRSINAOHTQRLRIDIHPPATLMGE LFGITGS
	2005..3144
	/gene="PRRL"
	/note="truncated receptor"
	/codon_start=1
	/product="nuclear hormone receptor PRRL-B"
	/protein_id="CAB55491.1"
	/db_xref="GI:5852065"
CDS	

/translation="MTCGCKGFFRRANKRNARLCPFRKAGEITKTRCOACRL
 RKCJESGKNNKIMSDAVERBALIRKRSERGTOLPVOGLTEORIMELIDA
 OKMTEDTFSHEKFRRLPGVLSGCELPESLOAPSPREARKNSOVRDLSLSVLOJ
 RGEQSYWNYKPPADSGSKETFLSLPHADWSTYTFGITSFKAVSTFDLEJEOI
 SLNGAEFLVQLQNFNTVFNATETWCGVRSCLDEPAGFQOLLEPMKRYHMK
 KLOJHEEYVLOMAISLESPDRPVLOHRYVDLOJEDFATLKSYTECNRPQPAHREL
 FLKIMAMITELRSINAOHTORLRIODIHPFAPIMOLFEGITS"

BASE COUNT 1157 a 1132 c 1095 g 1064 t
 ORIGIN

Query Match 92.5%; Score 2591.2; DB 52; Length 4448;
 Best Local Similarity 99.4%; Pred. No. 0;
 Matches 2632; Conservative 0; Mismatches 13; Indels 3; Gaps 3;

QY 139 agtggcttcaaacattcaaggagcccaagaagcaaaccttgaggtgagacccaaga 198
 DB 1802 ACTTACCAACCAAGTCCAGAGGCCAGAGCAAACTGAGGTGAGACCCAAAGAA 1861
 QY 199 gctgaacacatgctgactgttacctgtgaggaacacagagctgttccctggaagccca 258
 DB 1862 GCTGGAACCATATCTACTTGTGACACTGTAGGACACAGATCTGTCTTGGAAGCCCA 1921
 QY 259 gtgtcaacgcagatgaggaagtcggaaggtcccaaatctgcgtgtatgtgaggaacaag 318
 DB 1922 GTGTCAACGACAGATGAGGAAGTCCGAGAGTCCCAAAATCTGCCGTATGTGGGACAAAG 1961
 QY 319 ccaactggtatcaactcaatcatcatgatatgtgaagatgaggaaggttttcaaggagg 378
 DB 1982 CCACTGGCTATCACTCAATGATCATGATGAAAGTGAAGGCTTTTTCAGGAGG 2041
 QY 379 ccatagaacgcagacgcccgcgtgaggtgccccctccggaagggcgctgacagacc 438
 DB 2042 CCATGAACGCAAGCCGCGGTGAGGTGCCCTTCCGGAAGGGCGCTGCAAGATCAACC 2101
 QY 439 ggaagaccggtgacagatgacaggtcgctgcgctgcgaagtgctgaggaagcgcatga 498
 DB 2102 GGAAGACCCGCGACAGTGCAGCGCTGCCCTGCGCAAGTCCCTGAGAGCGCGCATGA 2161
 QY 499 agaagagatgatcatgtccgagagagccgtgaggaagggcgcttgataagagga 558
 DB 2162 AGAATGAGATATCATGTGTCAGACAGGCGGTGAGAGAGGGCGCTTGAATCAAGGGA 2221
 QY 559 agaaagatgaaacgcagaggtcagccactgaggtgacaggggtctgacagagagcagc 618
 DB 2222 AGAAAGTGAACGCAAGGAGCTAGCCACTGGAGTGCAGGGCTTACAGAGAGAGAGC 2281
 QY 619 ggaatgagatcaagagagatgagacgctcagatgaaacaccttgacataccttcc 678
 DB 2282 GGATGATGATCAGGAGAGCTGATGAGAGCTCAGATGAAACCTTTGACACTACCTTCC 2341
 QY 679 attcaagaatttcggctgacaggggtgcttagcagtgctgaggttgcagagctc 738
 DB 2342 ATTTCAGAAATTTCCGGCTGCCAGAGGGTCTTACAGATGGCTGCAAGAGCTC 2401
 QY 739 tgcagagcccatcagaggaagacgtgccagatgagacaggtccggaagagatctgtc 798
 DB 2402 TGCAGGCCCATCGAGGAAGAGCTGCAAGTGGAGCCAGGTCCGGAAGATCTGTGCT 2461
 QY 799 cttggaaggtctctctgacgtgcgagggagagatgaggtctggaactacaacccc 838
 DB 2462 CTTTGAAGGTCTCTCTCAGCTGCGGGGGGAGATGAGATGTCTGGAATACAAACCCC 2521
 QY 859 cagcgcagatgaggggaaagagatcttccctgctgcgcacacatgctggaatgtcaa 918
 DB 2522 CAGCCGACAGTGGCGGAAAGAGATCTTCTTCCTGCTGCCCCACATGGCTGACATGTCAA 2581
 QY 919 cctacatgttcaaaagcatcatcagcttgcacaaagtcatctcactactcaaggacttcg 978
 DB 2582 CCTACATGTTCAAAGGACATCATGCTTTCGCAAAAGTCACTCTACTTCAAGGAGACTGC 2641
 QY 979 ccaatcgagacagatccctctgtaagggggcgcttccagagctgtgtaactgaat 1038
 DB 3722 CAAATGTCAAGAGCTTGGCATGACCTCATTCGGGACCATCATCTGTGTGATCC 3781

DB 2642 CTATCGAGAGACAGATCTCCCTGCTGAAGGGGGCGCTTTCAGAGCTGATCACTGAGAT 2701
 QY 1039 tcaacacagtttcaacgcggagagatggaacctggagatgtgcccgtgtctactctc 1098
 DB 2702 TCAACACAGTTCACACGGGAGAGCTGGAACCTGGAGAGTGTGGCGGCTGTCTACTGCT 2761
 QY 1099 tgaagaacactgaggtgcttccagcaactctactggaagcccatgctgaattccact 1158
 DB 2762 TGAAGACATTCGAGAGTGGCTTCCAGCAACTTCTACTGAGGCCCATGCTGAATTCCTACT 2821
 QY 1159 acatgtgaagaagctgacgtgcatgaggaagatgltgtgtatgtgaagccatctcc 1218
 DB 2822 ACATGCTGAAGAGCTGACGTGATGATGAGAGAGATGTGCTGATGACAGGCACTCTCC 2881
 QY 1219 tcttcccccagacgcgcgaagttgtgtgacacccggtgtgtgtaacagctgagaagc 1278
 DB 2882 TCTTCCCCAGACCGCCGAGGTGTGACACACCGGTGTGTGACACAGCTGCGAGAGC 2941
 QY 1279 aattgcacatctctgaaagtctcaattgaattgaattgagtcggccagctgtcataagt 1338
 DB 2942 AATTGCCATTAATCTGAAAGTCTGAAAGTCTCAATTGAATGCAATGCGCCCAAGCTCTTAAGT 3001
 QY 1339 tctgttccatgaagatcatggtatgtctacccagatcccgacatcaatgtctcagaca 1398
 DB 3002 TCTTGTCTCTGAAGATCATGCTATGCTACACCGACTCCGAGCATCAATGCTCAGACA 3061
 QY 1399 cccagaggtgtgtgcacatccagaaataaaccttctgtcagccctcatgagaagt 1458
 DB 3062 CCAAGGCGTGTGCGCACACAGACATACACCCCTTGTCTACCGCCCTCATGCGAGGT 3121
 QY 1459 tfttgcgacacacagatgactgagcgctgcctt-aggatgaacctccggaaggaagc 1517
 DB 3122 TGTTCGACATCAACAGTACTGACGCGCTGCCCTTGGGGTGAACCTCCGAGAGAGAC 3181
 QY 1518 agacccaagccctctgagccgcaactcccggtccaagacagatgagacatgccaagagc 1577
 DB 3182 AGACCCAGAGCCCTGTGAGCCGCACTCCCGGAGCAAGATGAGACATGCCAAGAGC 3241
 QY 1578 ggaagaatgctgctggtcgtctctcctaaggaattctgtctatgacagctggtcagat 1637
 DB 3242 CGAATATGCTGTGCTGCTGTCTCTCTAGGAATTCCTGTATGACAGCTGGCTAGCAT 3301
 QY 1638 tccatgaagaagacatggtgtgcccccaaccccaagttcaagtctgaggaatgagacac 1697
 DB 3302 TCCATGAGAGAGACATGGGTGCCCCCAACCCCACTTGAAGTCTGAGGATGAAAGCCAC 3361
 QY 1698 agacttcaagtgtgagagtgacatgacctgtatgtcagagccatcaagaaggaagtgtg 1757
 DB 3362 AGACTCTTACGTGAGAGTGCATGACCTGTAGTCAAGACCATCAGAGAGCAAGGTTG 3421
 QY 1758 ccccttcccttaaaagccctgtgtgtgtggaagaaatccctaagatcccaataagt 1817
 DB 3422 CCGTTTCCCTTAAAGGCGCTGTGTGTGGGAGAAATCCCTCAGATCCCACTAAAGTG 3481
 QY 1818 tcaaggtgtgaaagagacaaagcgacaaagaaagaaagaaagaaagaaagaaagaaagaa 1877
 DB 3482 TCAAGGTGGAAGGAGACCAAGCGACCAAGATGAGCATGTGGGCTATAGCCACATA 3541
 QY 1878 cccaagttgtcgtctcctgagttcttcaattgtctacatcaatagctcgtctccac 1937
 DB 3542 CCAAGTTTGTGCTTCTCTGAGTCTTTCATTTGCTACTTAATAGCTGCTCCAC 3601
 QY 1938 ttcacacagttccctcctctctcagagctccttgggtgtaagacctgtatactcg 1997
 DB 3602 TTCACACTGTGCTCCCTCTCTCCAGACTCTCTTGTGGGCTCAAGGCTGTACTCATCG 3661
 QY 1998 gcaagtgatgagatctgtgtgagttcctcagaagaaagaaagaaagaaagaaagaaagaa 2057
 DB 3662 GCAAGTGATGATGATCTGTGGAGTCTCTTGAAGATGAGAAAGCAAGAGGCTGCTCAC 3721
 QY 2058 caaatgtcaagagcttggcatgacctcatccggacacatcatctgtctcgtaccc 2117
 DB 3722 CAAATGTCAAGAGCTTGGCATGACCTCATTCGGGACCATCATCTGTGTGATCC 3781

```

OY 2118 attgaacacataataagcactgaataagtagcctctgtgtgtgtatatacagcatga 2177
    |||||||
DB 3782 ATTGACACATATTATTAAGCACCGATTAATAGTAGCCTGTGTGGGGATATACAGCATGA 3841
OY 2178 cccagatataagctctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2236
    |||||||
DB 3842 CTCACATATATAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3901
OY 2237 acaatttgatcaaaagagaagaatgatgaatgaagaagaagaagaagaatttcctgt 2296
    |||||||
DB 3902 ACAATTTGGATCAAAAGAG -AATGATATAGTACACAAAGACGCAAGAAATTTCCCTG 3960
OY 2297 tctgaatctgaagctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2356
    |||||||
DB 3961 TGTGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4020
OY 2357 gactctgtgaagaagaagaagaagaagaagaagaagaagaagaagaagaagaagaaga 2416
    |||||||
DB 4021 GAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4080
OY 2417 gcttaggtctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2476
    |||||||
DB 4081 GGTAGGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4140
OY 2477 atgctctgtgaagaagaagaagaagaagaagaagaagaagaagaagaagaagaagaaga 2536
    |||||||
DB 4141 ATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4200
OY 2537 atgcaacctatattctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2596
    |||||||
DB 4201 ATGCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4260
OY 2597 gctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2656
    |||||||
DB 4261 GCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4320
OY 2657 tataaggaatccacacacacacacacacacacacacacacacacacacacacacacac 2716
    |||||||
DB 4321 TATAAGGATTCACACACCTAAGACTAGTTTGGGAAATGTAGCCCTGGTTTATGTCA 4380
OY 2717 aatcaaggaagaagaagaagaagaagaagaagaagaagaagaagaagaagaagaaga 2776
    |||||||
DB 4381 AATCAAGGCAAAAGAAATTAATATGTACTTTGGCTAAATAAAAAAAAAAAAAA 4440
OY 2777 aaaaaaaa 2784
    |||||||
DB 4441 AAAAAAAAA 4448

```

```

RESULT 4
LOCUS HSAJ9937 4337 bp mRNA PRI 09-SEP-1999
DEFINITION Homo sapiens mRNA for nuclear hormone receptor PRR2.
ACCESSION AJ0009937
VERSION AJ0009937.1 GI:5852066
KEYWORDS nuclear hormone receptor; orphan nuclear receptor; PRR2 gene.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 4337)
AUTHORS Heard,D.J., Holloway,J., Hansen,C., Tommerup,N., Aagaard,L. and
Vissing,H.
TITLE Identification of a novel protein isoform of the human nuclear
hormone receptor PXR/SXR and localization to chromosome 3q12.1-
13.3
JOURNAL Eur. J. Hum. Genet. In press
REFERENCE 2 (bases 1 to 4337)
AUTHORS Heard,D.J.
TITLE Direct Submission
JOURNAL Submitted (31-JUL-1998) Heard D.J., Novo Nordisk A/S, Molecular
Genetics, Novo Alle, DK-2880, Bagsvaerd, DENMARK

```

FEATURES

Location/Qualifiers

1..4337

source

Organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="3"

/map="q12-q13.3"

/tissue_type="liver"

/tissue_type="intestine"

/tissue_type="colon"

1771..3033

gene

/gene="PRR2"

1771..3033

CDS

/note="isoform with longer A/B domain, translation
initiates at GUG-1771 5 fold less efficiently than at
CUG-1840"

/codon_start=1

/trans_except="(pos:1771..1773,aa:Met)"

/product="nuclear hormone receptor PRR2-C"

/protein_id="CAB55493.1"

/db_xref="GI:5852068"

/translation="MDPRGEVGAKNLPNSPRGPANLEVRPKESMNTDFVHCEDTE
SVPRKPSVNADEVEGVPQICRYCGKATGYNHNVTCGCKGFPFRANKRNARLCP
RKGAEITKTRKQOCALRKCLSGMKNEKINSDAEVEERAILIKRKSERTGTOP
LGVQGLTEQRMVIRELMDAOKRTDITFSHKRNRSVLQJLGEDGSVWYKPPADSG
GREIFSLPHMADSTYMEKGIISPAKVISYRDPIEDQISLKGAEFLCOLRNT
VNAETGTEGRLSYCLEDTAGFQOOLLEBMLKFHMLKRLQHEEYVLMQAFSL
FSDPRGVLOHRYVDLOLOFAITLKSYTECNRPQPARFLFLKIMALTLELSINO
HRORLRIDIHFPATPLMQLFGITGS"

1840..3033

CDS

/gene="PRR2"

/function="transcription factor"

/note="full length receptor"

/codon_start=1

/product="nuclear hormone receptor PRR2-A"

/protein_id="CAB55492.1"

/db_xref="GI:5852067"

/translation="MEVPRKESMNTDFVHCEDTESVPRKPSVNADEVEGVPQICRYC
GKATGYNHNVTCGCKGFPFRANKRNARLCPRKGCEITKTRKQOCALRKCLSG
MKNEKINSDAEVEERAILIKRKSERTGTOPAGVGLTEQRMVIRELMDAOKRTD
ITFSHKRNRSVLQJLGEDGSVWYKPPADSGGREIFSLPHMADSTYMEKGIISPA
KVISYRDPIEDQISLKGAEFLCOLRNTVNAETGTEGRLSYCLEDTAGFQOOLLE
BMLKFHMLKRLQHEEYVLMQAFSLFSDPRGVLOHRYVDLOLOFAITLKSYTECN
RPQPARFLFLKIMALTLELSINAOHTORLRIIDIHFPATPLMQLFGITGS"

2005..3033

CDS

/gene="PRR2"

/note="truncated receptor"

/codon_start=1

/product="nuclear hormone receptor PRR2-B"

/protein_id="CAB55494.1"

/db_xref="GI:5852069"

/translation="MTCBGCKGFPFRANKRNARLCPRKGCEITKTRKQOCALRKCL
SGMKNEKINSDAEVEERAILIKRKSERTGTOPAGVGLTEQRMVIRELMDAOKRT
DITFSHKRNRSVLQJLGEDGSVWYKPPADSGGREIFSLPHMADSTYMEKGIISPA
KVISYRDPIEDQISLKGAEFLCOLRNTVNAETGTEGRLSYCLEDTAGFQOOLLE
BMLKFHMLKRLQHEEYVLMQAFSLFSDPRGVLOHRYVDLOLOFAITLKSYTECN
RPQPARFLFLKIMALTLELSINAOHTORLRIIDIHFPATPLMQLFGITGS"

2005..3033

BASE COUNT

1134 a 1105 c 1056 g 1042 t

ORIGIN

Query Match 84.2%; Score 2359.2; DB 52; Length 4337;

Best Local Similarity 95.2%; Pred. No. 0;

Matches 2521; Conservative 0; Mismatches 13; Indels 114; Gaps 4;

OY 139 agttgctcaaacaccacagagagcccaagaagaagaagaagaagaagaagaagaaga 198

DB 1802 ACTTACACCAACAGCCAAAGAGGCCAGAACCAACCTCGAGGCGTACAGACCAAGAAA 1861

OY 199 gctggaaccatgtgactgtgtacactgtgtgagacacagaagctgtgtctgtgaaagcca 258

|||||

Db	1862	GCTGGAAACCACTACGACTTGTGTACAGTGTGAGGACACAGAGTCTGTGTTCTGTGAAAGCCCA	1921
Qy	259	gtgtcaacgcagatgaggaagtcggaaggtcccaaatctgcggtgtatgtgtggacaag	318
Db	1922	GTGTCAACGCGAGATGAGGAAGTGGGAGGTCCCAAAATCTGCGGTGTATGTGGGGCAAGG	1981
Qy	319	ccatctgcctacacttaactgtcatatgcatgacatgtgaagatgtcaagagcttttccaagaag	378
Db	1982	CCACTGGCTATCACTTCAATGTCTATGTACATGTGTAAAGATGCAAGGGCTTTTTCAGAGAGG	2041
Qy	379	ccatgtaaacgcgaacgcgcgcgcctgtgagtgtccctctccgaaaggcgccctgtgcgaataccc	438
Db	2042	CCATGAAACGCAAGCGCCGGGTGAGGTGCCCTTCCGGAAAGGGCGCGTGGAGATCACCC	2101
Qy	439	ggaagaaccgcgcgcagctgcgcagcctgtccgcctcgcgaagtgtcctgtgaagcgcgtatga	498
Db	2102	GGAAGACCCGCGCAGATGTGCGAGCGCTCGCGCTCGCGCAATGTGCTGTGAGACGCGCAAGA	2161
Qy	499	agaagaagatgatcatatctccgaacgaaagccgttgaaagaaagagggcgcttgatccaagcgga	558
Db	2162	AGAAATGAGATATATATCTCCGACAGGCGCTGGAGGAGAGCGCGGCTTGATCAAGCGGA	2221
Qy	559	agaaagtgtgaacgcgacgcgagcctcaacgcaccttgagagtgacaaggcgctgcacaagagagcgc	618
Db	2222	AGAAAAGTGAACGACGACGAGGACTTCAGGCACCTGGGAGATGCAAGGGGCTGACAGAGAGCAGC	2281
Qy	619	ggaatgatgatcaaggagcgtgaatgagcgtcagaatgaataaccttgacatactctctcc	678
Db	2282	GGATGATGATATCAGGAGACTGATGGACGCTTCAGATGAACCTTTTGACACTACCTTCTCC	2341
Qy	679	attccaagaattctcgcgtctgcgaagggtgtctatgacgtgtgcgcgaattgtccaagatctc	738
Db	2342	ATTTCAGAAATTTCCG-----	2357
Qy	739	tgacgcgccatcgaaggaaagcgtgcacaagtgtgagcgacagtlccgaaagaatctgtgt	798
Db	2358	-----	2357
Qy	799	cttgaaggtctctctgcagcctgcgcggggagagatgacgtgtctgtgaatacacaacccc	858
Db	2358	-----GGTCTCTCTCACTCGCGGGGGAGGATGGCAGTGTCTGGAACCTCAAAACCC	2410
Qy	859	cagcgcgcagatgtgggggaaagaatatctctccctgtgtgcccaatgtgctgaacatgttaa	918
Db	2411	CAGCGGACAGTGGGGGAAAGAAATCTCTCCCTCGGCCCAACATGGCTGACATGTAA	2470
Qy	919	ccctcaatgttcaaaagagcatcatcagctgttgcgaagtcatctccacttccagggatctgc	978
Db	2471	CTCTATATGTTCAAAGGCAATATCAGCTTTGCCAAATCATCTCTCTACTTCAGAGGACTTGC	2530
Qy	979	ccatcgcaggaacagatctccctctgtgtgaaggggcgcttctgagcgtgtgtccaactgaagat	1038
Db	2531	CTATCGAGAGACAGATCTCCCTGCTGTAAGGGGGCGCTTTCGAGCTGTGTCAACTGAGAT	2590
Qy	1039	tcaacaacagtgttcaaacgcgcgagacatcggaacctgggaagtgtgtgcgcgctgtctctactgtc	1098
Db	2591	TCAACACAGTGTTCAAAGCGGAAACGTGGAACCTGGGAGAGTGGCGGCTGTCTACTCT	2650
Qy	1099	tggaaagcaactgcgcggtgtgtctccagaacattactgtgaagccaatgtcgtgaattccact	1158
Db	2651	TGGAGAGACATCGCAGGTGGCTTCCAGCAACTTCTATGTGAAGCCCATGCTGAATTTCCACT	2710
Qy	1159	acatgtctgaagaagctgcagcctgcgatgagagagagatgtgtgtgaatgcagcgcatactcc	1218
Db	2711	ACATGTCTGAAGAAGCTGCAGCTGGATGAGAGAGAGATGTGCTGATGCAAGGCGCAATCTCC	2770
Qy	1219	tctctctcccgacgcgcgaagtgtgtgtcagcaacgcggtgtgtgaaccaagcctgtcaagaagc	1278
Db	2771	TCTTCTCTCCCGACGCGCCGAGGTGTGTGTGTCAGACACGCGCTGTGTGACACAGCTGCAGGAGC	2830
Qy	1279	aattgcgcattactcttaaatctctaattgaatgaataatgtggcccaagcctgtcatagat	1338
Db	2831	AATTGCGCATTACTCTTAAGTCCCTACTTTGAATGAATGCAATGGGCCACAGCCTGTCTATAGT	2890

OY	1339	tcggttcctcgaagatcacaatgagctatgctatccagagctctccgaagatcaatgctcagaaca	1398
Db	2891	TCGTGTCCGGAAGATCATAGCTATGCTATCCGAGCTCCGAGATCATAACTCATGAGACA	2950
OY	1399	cccaagcgctcgtcgcgcataccagagacatacaacctcttgctacgcccctcatgcaagagt	1458
Db	2951	CCGAGGGGCGTGTGGCATCCAGAGCATATACACCCCTTTGCTAACGCCCTCATATGAGGAGT	3010
OY	1459	tgttcgcatcacaggtlaagctgaagcgctgcctt-aggctacacctccgaaggaagcgc	1517
Db	3011	TGTTGGGCAATCAACAGTAGCTAGTCAGCGGCTGCCCTTGGGGGTGACACCTCCGAGAGCGACC	3070
OY	1518	agaccagagcctctcgtgagcgcgcaactcccgaggccaacacagatggagacgtcccaagag	1577
Db	3071	AGACCCAGAGACCCCTGTAGAGCCGCACTCCCGGGCCAAACAGATGGACACTGCCAAAGGC	3130
OY	1578	cgacaatggccctcgtgcgcgtctccctctaggaattccctgtatgacagctggttagcat	1637
Db	3131	CGACAAAGCCCTGCTGCTGCTCTCCTCTAGGSAATTCCTGCTATGACAGCTGCGTACAT	3190
OY	1638	tcctcgaagaaagacatggtgtgccccccaccccccaagtlcaagtcgtlaaggaatgaagcac	1697
Db	3191	TCCTCAGAGAGGACATAGGGGTGCCCCCAACCCCACTTCAGTCGTAGAGGAGATGAAGCAC	3250
OY	1698	agactcttaagtggaagatgacactgctgtatgctcagagaccatcagaagaggaagttg	1757
Db	3251	AGACTCTTAAGTGTGAAGATGACACTGACCTGTGTGTCAGAGACATCAGAGAGGCAAGTTG	3310
OY	1758	cccttcctctttaaagggcccgctgtgtctcgtgggagaataccctcagaatcccaatgaagt	1817
Db	3311	CCCTTCTCTTTTAAAGGCCCTGTGTGTCTGSSGGAATAATCCCTCAGATCCCACTAAAGTG	3370
OY	1818	tcaagtggtgaaaggagccaagcgcagaccaagaaatgaagcatctggggtctatgcccacata	1877
Db	3371	TCAAGGTGTGAAGGGAGCCAGACAGACAAAGATAGGCCATCTGGGGGTCTATGCCACATA	3430
OY	1878	cccaagttgtctgcctctcctagatcttttcaatgctatcacttaatagtctgtctccac	1937
Db	3431	CCCAAGTTTGTGCTTCCTGCTTCGTGATCTTTTCATTCGTACTTATATAGCCGTCTCCAC	3490
OY	1938	ttcccaatgcttccctcctcctctccagagctgtttgtgggctaaagccgttactacgt	1997
Db	3491	TTCCCACTGCTTCCCTCCTCCTCTTCGAGCTGCTTTGTGGGCTCCAAGCCGTATCTATCG	3550
OY	1998	gcaagtgcatagatctgttggagtgctctctagagagaatgaagccaaggaagcgctgcac	2057
Db	3551	GCAGGTGCATAGATATCTGTGGGAGTCCCTATGAGATGAGAAAGCAAGAGGCGCTGCAC	3610
OY	2058	caaatgtcagaagccttgagcatgaactcaatccggccacatactctgtctcgtaccc	2117
Db	3611	CAAAATGCAGAACCTTGGCATGACCTCATTCGGCCACATCATTCGTGTCTGCATCC	3670
OY	2118	attggaacattattaaagcactgtaataaggtlagcgcgcgttgggtatadagcatga	2177
Db	3671	ATTGGAACAATTATTAAAGCAGCAGATTAATAGGTATACCTGCTGTGGGATATACACATTGA	3730
OY	2178	ctcagatatagatcccgagctcacagagtttlaagt-aaaaaaaacaacagaacaaca	2236
Db	3731	CTCGATATATAGATCCGAGCTCAGAGATTATTAATTAATAAAAAACAACGAACAACAA	3790
OY	2237	acaaattggttcaaaagsgaanaatgatatagtacaaaagcagacagaagaattccctg	2296
Db	3791	ACAAATTGGATCAAAAAGGAG-AAATGATATGACAAAAAGCAGCAACAAGAAATTTCCCTG	3849
OY	2297	tgtgtgatgctgagagcttgtatgacaggaactggtgtacccaagttaagttccccgaagcat	2356
Db	3850	TGTGTGATGCTGACCTGTGATGGGGGCACATGGGTACCCAACTGAAGGTCCCGAGGACAT	3909
OY	2357	gagctcttaagacaagggcacaaactgcaagctgtgagtgctgtgtgattggtgtga	2416
Db	3910	GAGTCTTATAGACAAAGGCGACAAACTGCAGCTGTGATGCGGTGTGTGATTTGGTGTA	3969

QY	2417	ggttaggtctgtcttccaccattatgaggcgcttggtttctctctggggcgctgagatgctgggt	2476
Db	3970	GGTAGGTCGTGTTGGCACTTATGAGGGCGCTGGATTGTCTCTGGGGCTGGAACTCTGGGT	4029
QY	2477	atgctctgtgcaaaaggcctgacgctgacaaatcagtttaaacacacccggaaagaaacatttac	2536
Db	4030	ATGCTCTGTGCAAAAGGCCTACCGCTGACATCAAGTTAAACACACCGGAGAAACATTTTAC	4089
QY	2537	atgcacctaatactctgtgtgtacacatctatctcctaaagcctaaaggatgaaagtcct	2596
Db	4090	ATGCACCTTATATTTCTGTGTACACATCTATTTCTTCACAAAGCTAAAGGGATGAAAGTGCT	4149
QY	2597	gacctgtttacagccactgtgagtgaaaaattttttgcatcttcacaaatatatactta	2656
Db	4150	GCCCTGTTTATAGCCACTGTGAGTAAAAAATTTTTCATCTTTCACAAATTAATCACTTA	4209
QY	2657	tataagcattccacacactaaagaaactagttttggaaatgtagccctggtttaatgtca	2716
Db	4210	TATAGGCAATTCACACCTTAAGAACTATTTGGAAATGTAAGCCCTGGTTTATATGTCA	4269
QY	2717	aatcaaggcaaaaggaaattaaataatgtactttggctaaaaaataaaaaaataaaaaa	2776
Db	4270	AATCAAGCAAAAGAAATTAATATGTAATCTTTGGCTAAAAAATAAAAAAATAAAAAA	4329
QY	2777	aaaaaaaaa 2784	
Db	4330	AAAAAAAAA 4337	
RESULT 5			
AF061056	LOCUS	AF061056 2146 bp mRNA	PRI 20-JAN-1999
DEFINITION	Homo sapiens orphan nuclear receptor PKR mRNA, complete cds.		
ACCESSION	AF061056		
VERSION	AF061056.1 GI:3511137		
KEYWORDS	human.		
SOURCE	ORGANISM	Homo sapiens	
		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
		Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.	
REFERENCE	AUTHORS	Lehmann, J.M., McKee, D.D., Watson, M.A., Willson, T.M., Moore, J.T. and Killewer, S.A.	
	TITLE	The human orphan nuclear receptor PKR is activated by compounds that regulate CYP3A4 gene expression and cause drug interactions	
	JOURNAL	J. Clin. Invest. 102 (5), 1016-1023 (1998)	
	MEDLINE	98395173	
	REFERENCE	2 (bases 1 to 2146)	
	AUTHORS	Lehmann, J.M., McKee, D.D., Watson, M.A., Willson, T.M., Moore, J.T. and Killewer, S.A.	
	TITLE	Direct Submission	
	JOURNAL	Submitted (21-APR-1998) Molecular Endocrinology, Glaxo Wellcome, 5 Moore Drive, RTP, NC 27709, USA	
FEATURES	SOURCE	Location/Qualifiers	
		1..2146	
		/organism="Homo sapiens"	
		/db_xref="taxon:9606"	
		/tissue_type="liver"	
		304..1608	
		/codon_start=1	
		/product="orphan nuclear receptor PKR"	
		/protein_id="AAD05436.1"	
		/db_xref="GI:3511138"	
		/translation="MEVPKPSNMHADPVHCEDESVGKPSVNADEVEGPOQICRVCA	
		GDKAGVFNWMTGCGKCGFRRARKRNARLRCPKRCACITRTTRBOCACLRK	
		LESGRKKEKINSDEAVERRRALIRKKEERGTQPLGQIGLEDRMMIRIMDAQNR	
		TFDTTSHFKNRLPGVILSSGCELPESIQAPSREAAASQVRDCLSKLSIDURGE	
		DGSVNNYKPPADSGGKEKESFLIPHADMSITWFKQISFPAVISYFRLPTEDISL	
		KGAAEELCOLRNTVFNAGETWEGCGRSLCYCEDPAGCFQOLLLEPMKPYMKIKKLC	
		LHEEYVIAQMAISLSPSPGVLOHRVYDLOEOPATIKSYIEENRQPAHRFLFLK	
		IMATTELRISNAQTORLLRIODIHPRATPLMQRLFOITGS"	
BASE COUNT	509 a	593 c	605 g 439 t
ORIGIN			

Query Match:	56.64:	Score 1866.6:	DB 10:	Length 2146:
Best Local Similarity	99.58:	Pred. No. 0:		
Matches 1812:	Conservative	0: Mismatches	9: Indels	0: Gaps
QY	139	agttggtctcaaacatcatccaagagagcccaagaaacaccttgaaagttagagaccacaagaa	198	
DB	266	acttaccacccaagcagctccaaagagcccaagaaacaaacttgaaagttagagaccacaagaa	325	
QY	199	gcttgaaaccaatgcctgaacttctgaacatgttgagacaagaatctgtctccctgaaagccca	238	
DB	326	gcctggaacacatctcaacttggatgacacttgagagacacagactctgttccgaaagccca	385	
QY	259	gtgtcaaacccaatgaatgaatgcgaaggttccccaatctgcgtaattgttgagagacaag	318	
DB	386	gtttcaccacccaatgaatgaatgcgaaggttccccaatctgcgtaattgttgagagacaag	445	
QY	319	ccactggtatactactcaatgtatcatatgaatgttgaaagatgcaagggctttttcaagagg	378	
DB	446	ccactggctatcatcttcaatgtatcatatgaatgttgaaagatgcaagggctttttcagagagg	505	
QY	379	ccattgaaagcaaacgcccggctgaaagtgcaccttcggaaaggggaccttgagagatacc	438	
DB	506	ccattgaaagcaaacgcccggctgaaagtgcaccttcggaaaggggaccttgagagatacc	565	
QY	439	ggaaagaccggcgagacatgttccagagcttcgcgccttgagcaagtgacctgagagcggtga	498	
DB	566	ggaaagaccggggagacatgttccagagcttcgcgccttcgcaagtgacctgagagcggtga	625	
QY	499	agaagaagatgatcatatgttcgaagagagccgttgagagagagcgggaccttgatcaagcga	558	
DB	626	agaagaagatgatcatatgttcgaagagagccgttgagagagagcgggaccttgatcaagcga	685	
QY	559	agaaagagtbaacgagacagagatctcaagccaatgttgaggttgcaagggcttgacaagagcag	618	
DB	686	agaaagagtbaacgagacagagatctcaagccaatgttgaggttgcaagggcttgacaagagcag	745	
QY	619	ggaatgatcatcagggagatctgatgagagctcagaatgaacaccttgacaataccttctcc	678	
DB	746	ggaatgatcatcagggagatctgatgagagctcagaacacacacacacacacacacacacac	805	
QY	679	atttcaagaatttcggcgctgcgcaaggggttgcttaagaatgttgagaggttgcaagagcttc	738	
DB	806	atttcaagaatttcggcgctgcgcaaggggttgcttaagaatgttgagaggttgcaagagcttc	865	
QY	739	tgaagagcccatcgaagagagaaagatctgcacaaatgttgagccaggttcggaagaagatctgt	798	
DB	866	tgcagagcccatcgaagagagaaagatctgcacaaatgttgagccaggttcggaagaagatctgt	925	
QY	799	ctttgaaggtctcttcgcgaagctgcgagggggagagatggaaggtgttcggaactcaaaaccc	858	
DB	926	ctttgaaggtctcttcgcgaagctgcgagggggagagatggaaggtgttcggaactcaaaaccc	985	
QY	859	cagccgaagcttgccgagaaagagatcttccctgcctgcacacaaatggctgacatgttcaa	918	
DB	986	cagccgaagcttgccgagaaagagatcttccctgcctgcacacaaatggctgacatgttcaa	1045	
QY	919	ccatcatgttcaagaagcatcatcagctcttgccaaagatcatctcctacttcagggacctgc	978	
DB	1046	ccatcatgttcaagaagcatcatcagctcttgccaaagatcatctcctacttcagggacctgc	1105	
QY	979	ccatcagggagcgaagatctccctgcctgaaagggggcgacttgagctgtgtcaaacatgagat	1038	
DB	1106	ccatcagggagcgaagatctccctgcctgaaagggggcgacttgagctgtgtcaaacatgagat	1165	
QY	1039	tcaaacagatgttcaaacgcgagaaatgtgaaccttgagagtggtggcggtgtctactgtct	1098	
DB	1166	tcaaacagatgttcaaacgcgagaaatgtgaaccttgagagtggtggcggtgtctactgtct	1235	
QY	1099	tggagagacatgtgaagtggtttccacgaacatcttacttgagagcccatgtctgaattccact	1158	
DB	1226	tggagagacatgtgaagtggtttccacgaacatcttacttgagagcccatgtctgaattccact	1285	

70788	70887	gap of unknown length
70888	77647	contig of 6760 bp in length
77648	77747	gap of unknown length
77748	83638	contig of 5891 bp in length
83639	83738	gap of unknown length
83739	88544	contig of 4806 bp in length
88545	88644	gap of unknown length
88645	91971	contig of 3327 bp in length
91972	92071	gap of unknown length
92072	98127	contig of 6056 bp in length
98128	98227	gap of unknown length
98228	104128	contig of 5901 bp in length
104129	104228	gap of unknown length
104229	107733	contig of 3504 bp in length
107733	107833	gap of unknown length
107833	112813	contig of 4982 bp in length
112813	112913	gap of unknown length
112913	118277	contig of 5363 bp in length
118278	118377	gap of unknown length
118378	122544	contig of 4167 bp in length
122545	122644	gap of unknown length
122645	125954	contig of 3310 bp in length
125955	126054	gap of unknown length
126055	129626	contig of 3572 bp in length
129627	129726	gap of unknown length
129727	132738	contig of 3012 bp in length
132739	132838	gap of unknown length
132839	135617	contig of 2779 bp in length
135618	135717	gap of unknown length
135718	138538	contig of 2821 bp in length
138539	138638	gap of unknown length
140793	140793	contig of 2155 bp in length
140794	140893	gap of unknown length
140894	142844	contig of 1951 bp in length
142845	142944	gap of unknown length
142945	145443	contig of 2499 bp in length
145444	145544	gap of unknown length
145544	147448	contig of 1905 bp in length
147449	147548	gap of unknown length
147549	150063	contig of 2515 bp in length
150064	150163	gap of unknown length
150164	151748	contig of 1585 bp in length
151749	151848	gap of unknown length
151849	153736	contig of 1888 bp in length
153737	153836	gap of unknown length
153837	154990	contig of 1154 bp in length
154991	155090	gap of unknown length
155091	156398	contig of 1308 bp in length
156399	156498	gap of unknown length
156499	157640	contig of 1142 bp in length
157641	157740	gap of unknown length
157741	159061	contig of 1321 bp in length
159062	159161	gap of unknown length
159162	160185	contig of 1024 bp in length.

FEATURES

source

1. 160185

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="3"

/clone="RP11-169N13"

45836 a 3206 c 31244 g 47693 t 3346 others

ORIGIN

	Query Match	Similarity	49.9%	Score 1398.8	DB 69	Length 160185
Best Local	Similarity	98.4%	Pred. No. 2.2e+282			
Matches 1423	Conservative	0	Mismatches 22	Indels 1	Gaps 1	
QY	1320	ccccgcgcgtccatcaggtctctgttccctggaagatcatgctcatgctacccgagctccgc	1379			
DB	68716	ctctggcttgacatgcagcttcttcttccctggaagatcatgctcatgctacccgagctccgc	68657			
QY	1380	agcatcaatgctcagcacacccagcgcgtctgcgcataccagacatacacccttgcct	1439			

Dh	68656	AGCATCAAGTCTAGACACACCAGGGCTGTGGCATCCAGGACATACACCCCTTGCT	685397
Qy	1440	acgcgcctcaatcgagagatgtgttcgcatcaagtgatcgaggggttcgcttggtga	1499
Dh	68596	ACGGCCCTCATGAGAGAGTGTTCGGCATCACAGGTACTGTAGGGCTGCTTGGGTGA	685337
Qy	1500	caactccgagagcgacccagaccgagagccctctgaagcccaatcccgggccaagaag	1559
Dh	68536	CACCTCCGAGAGGACACAGACCAGAGCCCTCTGAGCCGCCACTCCCGGGCCAAAGCA	68477
Qy	1560	atggaacatgcccgaagccggaacatgacctctgagctgtctccctagggaaatctgct	1619
Dh	68476	ATGGAACACTCCGAGAGCCGGAATGSCCTGTGCTGCTGCTCCCTAGGGAATTCCTGCT	68417
Qy	1620	atgacagctgggtatagcatctctcaaggaaagacatgggtgtcccccaaccccaatga	1679
Dh	68416	ATGACAGCTGGCTAGCATTCCTCAGAAAGAAATGGGTGCCCCACCCCACTCACTGCT	68357
Qy	1680	tgatggagatgaaagcacaagactctctaagtgaagatgacatgacccctgaatgaagac	1739
Dh	68356	TGATGGAGATGAAGCCACAGACTTTACGTGAGAGTCACTGACCTGTAGTACAGAAC	68297
Qy	1740	atcagagagcgaaagtgtgaccttcctcttaaaagccctgtgtgtctggtgagaaatccc	1799
Dh	68296	ATCAGAGAGCGAAGTGTGCCCTTCTCTTTAAAGGCCCTGTGCTGAGGAGAAATCCC	68237
Qy	1800	tcagatcccaataagtgatcaaggtgtggaaggaaccaagccaaggaatagcatct	1859
Dh	68236	TCAATATCCCACTAAATGTCTCAAGGTGTGGAAGGACCAAGGACCAAGGAAGGCCATCT	68177
Qy	1860	ggagtgcatgcccacacataccaaagttgtgtgctctccgagctcttcattgtcaacct	1919
Dh	68176	GGAGTCTATGGCCACATACCACAGTTGTGTCTTCTCTGAGTCTTTTCATGTGCACTCT	68117
Qy	1920	aatagtcctgtctcccaactcccaactcgctctccctctcttcacgagctgcttgtggct	1979
Dh	68116	AATAGTCTGTCTCCACTCTCCACTCGTTCCTCCCTCTCCAGGCTGCTTGTGGCT	68057
Qy	1980	caagggctgtactcaatcgcaaggtgcataagatctgtgggaagtcctctagaagaatgag	2039
Dh	68056	CAAGGCTGTACTCATATGGGACGGCCGATAGTCTGTGGGAGTCTCTTAAGAGAAAGAG	67997
Qy	2040	aagcagagagggccgcgcacaaatgctagaagaagtttgcatatgacctatccggacacat	2099
Dh	67996	AAGCAGAGAGGCTCGCACCAAAATGTAGAAAGCTTGCCATGACTCATATTCGGCCACTCA	67937
Qy	2100	tctcgtgtctcgtcatccatctgaacacatttaagcaactatataagtagaacctgtg	2159
Dh	67936	TTCGTGTCTCTGATCCATTTAACAACATTTTAAGCACGATTAATGTAAGGCTGCTG	67877
Qy	2160	tgagggtatcacgcatctgactcagataatagaccctgagctcacagaagttatagtaaaaa	2219
Dh	67876	TGGGGTATACAGCAATTGACTCAATATTAATACCTGAGCTCACAGAGTTTAATGATTA	67817
Qy	2220	aacaaatcagaanaaccacaacatttgatacaaaaagagaagaatgataatgacaaaagag	2279
Dh	67816	AACAAACAGAAACCAACACATTGTGATCAAAAGAG- AATATGATTAATGTACAAAAGAG	67757
Qy	2280	caacaaggaattccctgtgtggaatgctgagcctgtgaatgacagcaactggtgaaccaag	2339
Dh	67757	CACAAGGAATTTCCCTGTGTGATGCTGAGCTGTATGCGGGGCACTGGTAACCAAGTG	67697
Qy	2340	aaggttcccgagacatgagtcgtgtagagcaaggacaaaactgacgtgtgagtgctgt	2399
Dh	67697	AAGGTTCCTCAAGACATGATCTGTAGAGCAAGGCAAAACTGCACTGTGATGTGCTGT	67637
Qy	2400	gtgtgtatgttggtgtgtagtaagtcgtgttgcaactgtgtggagctggtgtgttctgt	2459
Dh	67637	GTGTGTATTTGTGTAGAGGTCTGTGTCCTACTTGAATGGGCTGTGGTTGTCTGTG	67577
Qy	2460	gggtcgagatctgtgagatgctctctgtgacaaggtacgctgacaaatcagtttaaacacac	2519

Dn	1004	TCAACGAGATCTTGACACCGAAGGAACCTGGAGATGGCGCGGTGACTTACAGCT	1063
Qy	1099	tgaagacactcgcggttgcttccagaacaattctaagtggaccatgctgaattccaact	1158
Dn	1064	TCGAAGAACCCTAAATGGTGGCTTCACAGAAACTTCGTGGATGCATTGATGAATAATCCACT	1123
Qy	1159	acatagcgaagaagactgtaagcttcgatctggagggttatgctcatatgaaagcaattccc	1218
Dn	1124	GCAATGCTGMAAACCTATACAGTGCATTAAGAGAGATATGCTCTATATGCAGGCCATCTCCC	1183
Qy	1219	tctctccccagagcccgaggtgtctgtcagaacccgctgtgtgtagaccagctcagaagc	1278
Dn	1184	TCTTCTCCCCAGATGCTCTGTGTGTGTGCTCAGACCGAGCTGGTAGACCAACTGCAGAGTA	1243
Qy	1279	aattgccattactctgaagctcctacattgaatgatcaatcggccccagctcctaaggt	1338
Dn	1244	GGTTTGCCCTCACCTCGAAGGCGCTCATATTGATGATGAGTCGGCCATATCTCTGCTCACAGT	1303
Qy	1339	tctgttctctgaagaatcatactgctatgctcacccagagctccgacagcatcaatgctcagacca	1398
Dn	1304	TCTCTTCTTCTTAAGATCTATGGCCGCTCTCTACTGATGACTCGAAGCATCAACGCCCAAAA	1363
Qy	1399	cgcagagcctgctgtgcacatccagaagacataccccctttgtcagagccctcatcagaagat	1458
Dn	1364	CCGACCACTTGCTGCGCATCCCAAGACTGCGACCCCTTTGCGCACCCCATCTCATGCAAGAT	1423
Qy	1459	tgttcgcatcacaggtlactgtcagagcgtctgaccttgggttgaaacctccgagagagacca	1518
Dn	1424	TATTATGACAGCAGCATGTGGCTGATGGGTGCGCCCTGATGAGGAGATCTCATGGGCGACGTA	1483
Qy	1519	gaaccagagccctctgagccgcgacctcccgggccaagacagatgtgacactgcgcaagagcc	1578
Dn	1484	GACCCAGA-CATTTCGAAATTCGCACTTCTAGGGCTAGACAGATGAGAC-TACTGATAGCC	1541
Qy	1579	gacaaatgcctc 1589	
Dn	1542	AACAACGCCCT 1552	
RESULT	10		
LOCUS	AF151377	1755 bp	mRNA ROD 05-AUG-1999
DEFINITION	Rattus norvegicus pregnane X receptor mRNA, complete cds.		
ACCESSION	AF151377		
VERSION	AF151377.1	GI:5702232	
KEYWORDS	.		
SOURCE	Norway rat.		
ORGANISM	Rattus norvegicus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;		
TITLE	Rattus.		
JOURNAL	1 (bases 1 to 1755)		
PUBMED	Zhang,H., Leculuyse,E., Liu,L., Hu,M., Matoney,L., Zhu,W. and Yan,B		
REFERENCE	Rat Pregnane X Receptor: Molecular Cloning, Tissue Distribution,		
AUTHORS	and Xenobiotic Regulation		
TITLE	Archive of Biochemistry and Biophysics 368 (1), 14-22 (1999)		
JOURNAL	10415106		
FEATURES	2 (bases 1 to 1755)		
source	Zhang,H., Leculuyse,E., Liu,L., Hu,M., Matoney,L., Zhu,W. and Yan,B		
	Direct Submission		
	Submitted (13-MAY-1999) Biomedical Sciences, University of Rhode		
	Island, 41 Lower College Road, Kingston, RI 02881, USA		
	Location/Qualifiers		
	1..1755		
	/organism="Rattus norvegicus"		
	/strain="Sprague-Dawley"		
	/db_xref="taxon:10116"		
	218..1513		
	/codon_start=1		
	/product="pregnane X receptor"		
	/protein_id="AAD47214.1"		
	/db_xref="GI:5702232"		
CDS			

Query Match	34.5%	Score 967.8	DB 11	Length 1755
Best Local Similarity	80.1%	Pred. No. 2,76-192		
Matches 1150	Conservative 0	Mismatches 282	Indels 3	Gaps 1
DB	458 a	468 c	480 g	349 t
ORIGIN				
Query Match				
139	agttggtcttaaacatccatccaagagagcccaagaacaaacttggaggtyagagcccaagaanaa	198		
174	AGTGGCACCTTACACTACACTGACAGACACACAGATGTTAAACCTGGAGATGAGACTGACAGAGA	233		
199	gctggaacccatctgtgactttgttaacactgtgtggagaaacagagctgtctccggaaagccca	258		
234	GGTGAACCTAGTGTGGCCCTTGATCAACGGGAAGACGATTCGTGTTGGAAGAGCC--	291		
259	gtgtcaacgacagatctgggaagtcggagagtcgcccaaatctgcgtgtatctgtgggagacaag	318		
292	-TATCAACGATGATGAGAGAAATGAGAGCTCTTCAAAATGCGCGTATGTGGGACACAG	350		
319	ccactgtgtctacatctcaatlytcatatgacatgtgaagagatgcaagagcttttcagaagag	378		
351	CCAATGGCTATCACTTCATATGTCATGATACCTGTGAAGATGTAAAGGATTTTTCAGAAAGG	410		
378	ccatgaagaacgacagcccggtctgaggtgtgcccttccggaaagagcgcccttgagagctaccc	438		
411	CCATGAAGCCCAATGACCCGCTGAGGTGCCCTTCCTCCGAAGGGAGCCTGCGAGATCACCC	470		
439	ggaagaccgagcagacagttgcagagcctctgcgcctctgcgaagttccctggagagcgcgatga	498		
471	GGAAAGACCCAGCGCACTGCTCCAGAGCTTGCCTTTTCCGAATGCTCGAAGAGTGGCATGA	530		
499	agaagagagatgatactgtccgaacgagcggtggaagagagagcgagccttgataaagcgga	558		
531	AGAAAGAGATGATCATGTGATGTGAGCGCGGTGTGAACAGAGGGGGCCTTGATCAAGAGGA	590		
559	agaaagatgaaacgagagatctcagcaactctgaggtgacaggggcttgacagaagagacagc	618		
591	AGAAAGAGGAAAGATTGAGGCTCCACCCCTGGAGGGAGGGGCTGCACAAACAGAC	650		
619	ggaatgatatcagagagatctatgagagcctcagaatgaaacatttgaacatactcttccccc	678		
651	AGGGCTGATCAAGAGACTGATGAGAGCTCAATGACCAACCTTTTACCAAACTTTTCTCC	710		
679	attcaagaatattccgagctgcacaggggtgcttagcaagtgagctgcaggttgccagaagcttc	738		
711	ACTTCAAGAGATTCCGGCTACCTGCGGTGTTCACAGTGAATGACGAGCTTCCGGAGGATTC	770		
739	tgcagagcccatcgaggagaaagactctgccaagtggagccaggttcgcgaagaatctgtgct	798		
771	TGCAGGCTCTACTGTTGGAAGACCTGCGCACATGAGTCAATTCATGAAACACAGTGTTC	830		
799	cttgaagatctctctcagactctcgagggggagagatgacagtgcttgagaaactacaacccc	858		
831	CAATCAAAATTTCTGTCTACACTCTCGGGAGAGACGGCAGCATCTGGAACTTACCAACCTC	890		
859	cagccgacagcggcgagaaagagatcttccctctgtgcgccacataggctgacatgtcaa	918		
891	CCTCCAAAGGAGGAGGAAAGATATATCCCTCTCTCGGAGACTGGCCGATGTGTCAA	950		
919	ctcacaatgttcaaaagacatcaagctttgccaagtcatctctactactcaggagactgtc	978		
951	CCTACATTTTCAAGGGGCTCATCAACTTGCACAAATCATATCCACATTCAGGAGGCTGC	1010		
979	ccatcgagagacagatctccctctgtaaggggcgcgcttgcagtgatgtgtcaactgagat	1038		


```

Db 1011 CTATGAGAGACGATCTCCCTGCTGAGAGGGGCCCATTTGAGATGTCATCCTGAGGT 1070
Qy 1039 tcaacacagtggttcaacagcggagacttgaaaccttgagtgagcgggtgtgcttactgct 1098
Db 1071 TCACACACATGTTTCACACGGAACACGAGACCTGGAGTGGGTGGCTGAGTACTGCT 1130
Qy 1099 tgaagacactgcaggttgcttccagcaacttctacttgagcccatgctgtaattccact 1158
Db 1131 TCAGAGACCTTAATAGGGGCTTCCAGAGCTCCCTGCTGAGACCCGTTGATGAAATTCAC 1190
Qy 1159 acatgctaaagagctgagctgacgtgagagagagatgtgtcgtatgcagagccatccc 1218
Db 1191 GCATGCTGAAGAGCTACAGCTGCTGAGAGAGTACGTCTGATGATGAGCCATCTCCC 1250
Qy 1219 tcttcccccagaccgcccaggtgtgtcgcagaccgcccgtgtgagccagctgagagac 1278
Db 1251 TCTTCTCCCGAGATCGCCCTGCTGCTGCTTCAACGTAAGCGTGTGAGACCTGCGAGAG 1310
Qy 1279 aattgcacattactctgaagtcctacattgaatgcaatcggcccccagcctgctcatagt 1338
Db 1311 GATTGGCCCTCACCCTGAAGGCTTACATCGAGTGAAGCGGCCCTTATCCTGCAACAGAGT 1370
Qy 1339 tctgttctgaagatatagtctatgtctacacagagctccgagacataatgtctcagcaca 1398
Db 1371 TCTGTCTCTGAAGATCATGGCTGCTCTCACACGAGCTGCGAGATGATGATGCCACAGAG 1430
Qy 1399 cccagcgcgtctgctgcacatcacagacatcacaccccttctgtagccctctatgcaagagt 1458
Db 1431 CCGAGAGACTACTGCGATCCAGAGACAGACGCCCTTGCACACCTCTATCATGAGAGAT 1490
Qy 1459 tgttcgcacatcacaggtagctgagcgcgtcgccttgggtgacacctccgagagagcaca 1518
Db 1491 TATTGAGAGAGCAGCAGAGGCTGAGTGGCTGCTGCTGAGTGAAGATCTCACGAGAGACCA 1550
Qy 1519 gaccacagaccccttgagcgcgcacccctccgagcacaagacagatgagacatgcacaa 1573
Db 1551 GACCCAGATGTTGAAATTCGCACTTCTAGGGCTATCAGATGACACACATGATTA 1605

```

```

RESULT 11
LOCUS 365 bp mRNA 01-JAN-1998
DEFINITION SHGC-56597 Human Homo sapiens STS cDNA, sequence tagged site.
ACCESSION G36929
VERSION G36929.1 GI:2734596
KEYWORDS STS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
AUTHORS Myers, R.M.
TITLE Human STS (1997)
JOURNAL Unpublished (1997)
COMMENT
Contact: Richard M. Myers
Stanford Human Genome Center (SHGC)
Stanford University School of Medicine
Department of Genetics, M-344, Stanford, CA 94305, USA
Tel: 4157259687
Fax: 4157259689
Email: myers@shgc.stanford.edu
Primer A: CCCATCAGTGGCAACAGCA
Primer B: AGTGAAGGTTCCCGAGGACA
STS size: 107
PCR Profile:
Initial incubation: 95 degrees C for 10 minutes
Denaturation: 94 degrees C for 30 seconds
Annealing: 60 degrees C for 30 seconds
Polymerization: 72 degrees C for 23 seconds
PCR Cycles: 30

```

```

Thermal Cycler: Perkin Elmer 9600
Protocol:
Template: 25 ng
Primer: each 1 uM
dNTPs: each 200 uM
Amplify: Gold Polymerase: 0.07 units/uL
Total Vol: 5 uL
Buffer:
-MgCl2: 2.5 mM
-KCl: 50 mM
-Tris-HCl: 10 mM
pH: 8.3
Prepared with primer pairs derived from N69311 -- Unigene.
FEATURES
source
1..365
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="3"
/clone_id="Human"
STS
Primer_bind 173..192
Primer_bind 173..192
BASE COUNT 95 a 106 c 68 g 95 t 1 others
ORIGIN
Query Match 12.0%; Score 335; DB 77; Length 365;
Best Local Similarity 97.0%; Pred. No. 6,6e-60;
Matches 352; Conservative 0; Mismatches 10; Indels 1; Gaps 1;

```

```

Qy 2251 aaggaagaatataagagacaaagc-agcacaagagattccctgtgtgagctgag 2309
Db 365 AAGAGGAATATATAGTACGACAAACACACAGAAATTCCTGTGTGATGCTGAG 306
Qy 2310 ctgtatgagcagcagctggtgacccaagtgaaggttcccgagacatgagctgtagag 2369
Db 305 CTGTATGCGCGGCGACTGGGTACCAAGTGAAGGTTCCCGAGACATGAGTCTGAGAG 246
Qy 2370 caagggcacaactgcaactgtgagtggtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2429
Db 245 CAAGGACACAACTGCACTGTGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 186
Qy 2430 gccacttgatggcctggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2489
Db 185 GCCACTGTATGAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 126
Qy 2490 aggtacagctgacaaatcaatgtaaacacacccgagagaaacattacatgacattat 2549
Db 125 AAGTACCTTACATCATGATTAAACACACCGAGAAAGACCATTTACATGACCTATAT 66
Qy 2550 ttctgtgacatcatcttccaaagcctaagaggtatgaagtgccctgtgtgtgtgt 2609
Db 65 TTCTGTGACATCATTTCTAAAGGCTAAAGGATGAAGTGCCTGCTGTTATAG 6
Qy 2610 cca 2612
Db 5 CCA 3

```

```

RESULT 12
LOCUS AF276753 1340 bp mRNA VRT 03-OCT-2000
DEFINITION Gallus gallus xenobiotic receptor (CXR) mRNA, complete cds.
ACCESSION AF276753
VERSION AF276753.1 GI:10505178
KEYWORDS
SOURCE chicken.
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.

```


GSNYSRPRPTLSFGNSSSSSDLYTSLDMESGFSNLDNEDDDPSVTLDS
 PLSMPLHNDLVSTSLQKIVGRKMKIPGRDLTSDQVLLKSSAIEVTKNSQST
 MODMSWDGSDQYKIDVTYVSGAGHLELLEPLIFQVGLKMLHEEHYLLNACI
 VSDRPGVDANKLVEALIDRLSLNLTQYTRCHRPSPGSHQVAKKIOALDRLSNEE
 HSKOYRSLSFOPENSKMLTPIVLEFGNHS"
 BASE COUNT 448 a 647 c 518 g 430 t
 ORIGIN Unreported.

Query Match 10.0%; Score 278.8; DB 11; Length 2043;
 Best Local Similarity 55.6%; Pred. No. 4.2e-48;
 Matches 629; Conservative 0; Mismatches 482; Indels 21; Gaps 4;

QY 288 ccccaaatccgctgtaagtgtgggagacagcactggctatacttcaatgtcatgaca 347
 DB 155 ccccgagctgctgtagagctgtgagacccgagccagcgttccatctaatgtatgacc 214
 QY 348 tctgag 407
 DB 215 tctgag 274
 QY 408 cccctccgag 467
 DB 275 ccccttc---aattgagatttcccgatccacagagacacgagacacgagacgcttcg 331
 QY 468 cgcctcgag 527
 DB 332 cgcctcamaagcgtgtag 391
 QY 528 gtggag 587
 DB 392 gtgag 451
 QY 588 ctggag 647
 DB 452 ctgag---gcccaagctatctgagagagagagagagagagagagagagagagagagag 508
 QY 648 cagatgag 702
 DB 509 caccacamaagcgtgtag 568
 QY 703 -gggtctgag 761
 DB 569 atggag 628
 QY 762 gctgcag 821
 DB 629 tcttccctcag 688
 QY 822 cgggag 881
 DB 689 tccaaagctgag 748
 QY 882 atcttctcctgctgag 941
 DB 749 cctctcctcag 808
 QY 942 agcttgcag 1001
 DB 809 ggccttgcag 868
 QY 1002 ctgag 1061
 DB 869 ctgag 928
 QY 1062 acgag 1112
 DB 929 gatag 988
 QY 1113 ggtgag 1172
 DB 989 gctggag 1048

QY 1173 ctgag 1332
 DB 1049 ctgag 1108
 QY 1233 cggcag 1292
 DB 1109 cgagctggggttcag 1168
 QY 1293 ctgag 1352
 DB 1169 ctgag 1228
 QY 1333 atcatgag 1404
 DB 1229 atgag 1280

RESULT 15
 AF011356
 LOCUS AF011356 2494 bp mRNA VRT 08-JUL-1997
 DEFINITION Gallus gallus vitamin D receptor mRNA, complete cds.
 ACCESSION AF011356
 VERSION AF011356.1 GI:2245698
 KEYWORDS
 SOURCE
 ORGANISM
 chicken.
 Gallus gallus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
 Phasianinae; Gallus.
 REFERENCE
 1 (bases 1 to 2494)
 Lu, Z., Hanson, K., and Deluca, H.F.
 Cloning and origin of the two forms of chicken vitamin D receptor
 Arch. Biochem. Biophys. 339 (1), 99-106 (1997)
 MEDLINE
 97223369
 JOHNNAL
 TITLE
 Direct Submision
 Lu, Z., Hanson, K., and Deluca, H.F.
 Submitted (30-JUN-1997) Biochemistry, University of
 Wisconsin-Madison, 420 Henry Mall, Madison, WI 53706, USA
 FEATURES
 source
 location/Qualifiers
 1. 2494
 /organism="Gallus gallus"
 /strain="leghorn cockerel"
 /db_xref="taxon:9031"
 /tissue_type="kidney"
 112. 1467
 /codon_start=1
 /product="vitamin D receptor"
 /protein_id="AAB62579.1"
 /db_xref="GI:2245698"
 /translation="MSEIRGSMDEQOQSNAYLPDADMDVVAASTSLPDAQDPDRVY
 RIGVCGDPRATGPHYANMCEGCKGFRSMKRAKFTPEPNDCITTDNRHRCAC
 RLKRCYDIGNAKKEFILTDERVQKREMLIKKEELKSLKPKLSEEDQKVIDTLE
 AHRKTDYTSDEKRPVPRKFSRMAHSSVVOGSPSSDSNDVGSFAFAFP
 EPEPEQMFSLNLDSESDSPSNMIEPLPHLPHLPHLADLVSYIQVIGFAKIPGR
 DLAEQDIALKSSALEVILKNSQSTFEDMSKSGNDFKYKVDVQAGSHMDL
 EPLVROVQKTLNHEEHVILMAICILISPDGQVDVSLVSPDRLSLDLYRIR
 GRPPGSRLLYKMKIQLADLSLNEHSKQIRCLSPQPHSMQTLPIVLEFGNET
 S"

CDS
 154. 1467
 /note="form B: coding region of second protein form
 translated from this single transcript because of 'leaky
 scanning' mechanism"
 BASE COUNT 536 a 781 c 745 g 432 t
 ORIGIN

Query Match 9.8%; Score 275; DB 4; Length 2494;
 Best Local Similarity 54.9%; Pred. No. 2.7e-47;
 Matches 670; Conservative 0; Mismatches 515; Indels 36; Gaps 5;

Db 241 CCCGATCTGCGGGGCTGCGGGGACAGACACCGGCTTCACATTCACAGCCATGACC 300
Qy 348 tctgaagatgcaaggcttttcaaggagccatgaaacgcaacgcccgtatgagtc 407
Db 301 TCGGAGGGCTGCAAGGGCTTCTTCAGGAGGAGCATGAAGAGAAAGGATTTCACTGT 360
Qy 408 ccccttcggaaggcgccctgcagatcaccccggaagaccgcgcagctgcagccctgc 467
Db 361 CCGTTCT---AACGGGAGACTGCAGAAATCACAAGAGCAACCGCGGCACTGCGAGCCTGC 417
Qy 468 cgcctgcgcaagtgcctgcagagcgcaatgaagaagagatgcatgctccgacgagcc 527
Db 418 CCGCTGAGAGCGCTGCGATTCGATGATGAAGAGTTCAACCTGACGAGAGAG 477
Qy 528 gtgagggagagcgcgcttatacgaagcgaaagaaatgaacgagacagagactcagca 587
Db 478 GTGAGAGAGAGCGCTGAGATGATCTGAAGCGCAAGAGAGAGAGCGCTGAAGAGAGC 537
Qy 588 ctgagagtgcaaggcgctacagagagagcgagtgatgacagtgatgagcgt 647
Db 538 CTG---AAGCCCAACTCTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGC 594
Qy 648 cagatgaaacaccttgacactacccctccattcaagaattccggctgcaggggtg 707
Db 595 CACGACAAAGCTTCGACACACCATCTCGACTTCAACAAGTTTCGGCCCGCGTGAGA 654
Qy 708 cttagcagtgctgcagatgcca-----gagctctcagggccca 749
Db 655 AGCAATTCAGAGAGAGATGGCAAGCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 714
Qy 750 tctgaaggga---agaagctgcgaagtgcagagcgagtcggaagaatctgtctcttgaag 806
Db 715 TCGGAGAGACTCCAGAGATGCTTTGGCTCCGATGCTCGCTCGATTCAGAGAGCCATG 774
Qy 807 gtctctgcagctgcgaggggagagatgpcagtgtctggaactaacaaccccgacgac 866
Db 775 GAGCCCCAGATGTTTCCAACTCGACCTCTCCGAGAGAGCGATGAGAGCCCTCCATG 834
Qy 867 agtgcgaggaagagatcttctccctgtgcccacatgagtgatgacatgtaactatg 926
Db 835 AACATCGAGCTGCCCCACCTCCCATGCTCCCGACCTGCGGACCTGTGTCACTACAGC 894
Qy 927 ttcaagagcatcatcagcttgcgaagatcatctcacttcaaggagctgcacatcgag 986
Db 895 ATACGAGAGATGATGCGCTTGCCCAAGATGATCCCGGATTCAGGAGATCTGACGGAGAG 954
Qy 987 gaccagatccctgcgtgcagggggcgcttgcagctgtgcaactgagatccaaca 1046
Db 955 GATCAGATTGCTCTGTAATCCAGCGCATCGAGGTGATCATGCTGCTCCAAACGAG 1014
Qy 1047 gtgttcaacgaggaactgcggaacctggaggtg-----tgccgctgttctactgc 1097
Db 1015 TCCTTCACCATGAGAGACATGCTCTGAGACTGCGGAGCAATGACTTCAAGTACAAATC 1074
Qy 1098 ttggaagacactgcagtgcttccagcaacttctactgcagcccatgtgaaattccac 1157
Db 1075 AGCGATGTCACCCAGCGGAGACAGAGATGAGACTCTCGGAGCGGCTCGTGAAGTTCAG 1134
Qy 1158 taatgctgaaagagctgcagctgcataggaagagatgtgtcgtatgcagggccatctcc 1217
Db 1135 GTGGGCTTAAAGAACTGAACCTTCACGAGAAAGAGCACTGCTGCTCATGCGCATCTGC 1194
Qy 1218 ctcttctcccgagacggccaggtgtgtgcagacggcggtgtgacacagctgcaagag 1277
Db 1195 ATCTGTCCCGAGATCGGCGCGCTTCAGAGACACTCTCGTGTGAGTTCATTCAGAGAC 1254
Qy 1278 caattcgcaattactctgaagtcctacattgaatcgaaatcgcccaagcctgtcatag 1337
Db 1255 CGGCTCTGGAGATCTCTCAGACCTACATCGCTGCAAGAGACCCCGCGGAGAGCGC 1314
Qy 1338 ttctgttccctgaagatcatgtgtatgtcacagagctccgagcatcatgtctcagcag 1397
Db 1315 CTGCTGTAGCGCAAGATGATCCAGAAAGCTGCGGAGCTGCGGAGCTCAAGAGAGAGCAGC 1374

Qy 1398 acccaagcgccgtgtgcacatccagagacatacaacccttgcagccctcatatgagag 1457
Db 1375 TCCAGAGATGACGGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTG 1434
Qy 1458 ttgttcgcatcagagtagc 1478
Db 1435 GTGCTGAGGTCTTCGCAAC 1455

Search completed: February 18, 2001, 16:48:29
Job time: 11089 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 18, 2001, 13:05:04 ; Search time: 67.2 seconds

(without alignments)
824.990 Million cell updates/sec

Title: US-09-143-828-4

Perfect score: 2494
Sequence: 1 MVTFRHHEKESGLRAPALP.....QDIHPFATPLMDELFGITGS 473

Scoring table: BIOSUM62
Gapop 10.0, Gapext 0.5

Searched: 374700 seqs, 117207915 residues

Total number of hits satisfying chosen parameters: 374700

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database:

SPTREMBL_15:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_protent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2494	100.0	473	4	Q9UNW4
2	2330	93.4	457	4	Q9UJ26
3	2278	91.3	434	4	Q9UJ27
4	2110.5	84.6	420	4	Q9UJ23
5	2058.5	82.5	397	4	Q9UJ24
6	1973	79.1	379	4	Q9UJ25
7	1814.5	72.8	411	6	Q9UJ02
8	1761.5	70.6	431	11	Q9RI17
9	1753.5	70.3	442	4	Q9UJ22
10	991.5	39.8	386	13	Q91839
11	817	32.8	420	13	Q91B73
12	796	31.9	453	13	Q9PTN2
13	790.5	31.7	425	13	Q91B74
14	729	29.2	358	11	Q9UJ51
15	504.5	20.2	445	11	Q9UJH7
16	500.5	20.1	455	4	Q12970
17	472.5	18.9	472	4	Q92843
18	463	18.6	469	5	Q92735
19	460	18.4	680	5	Q903Y4

20	459	18.4	397	5	Q76246	Q76246 uca pugliat
21	457	18.3	386	13	Q9PVE4	Q9PVE4 brachydanio
22	453	18.2	673	5	Q76827	Q76827 ceratilis c
23	452.5	18.1	484	11	Q60641	Q60641 mus musculu
24	450	18.0	416	13	Q9W6N4	Q9W6N4 hippoglossu
25	447.5	17.9	456	4	Q13986	Q13986 homo sapien
26	446	17.9	454	13	Q9W524	Q9W524 fugu rubrip
27	445	17.8	541	5	Q97095	Q97095 locusta mig
28	444	17.8	447	13	Q9W523	Q9W523 fugu rubrip
29	437	17.5	881	5	Q9V9K8	Q9V9K8 drosophila
30	433.5	17.4	457	13	Q90272	Q90272 brachydanio
31	432	17.3	444	13	Q91391	Q91391 brachydanio
32	431	17.3	491	5	Q02035	Q02035 tenebrio mo
33	428	17.2	454	13	Q90271	Q90271 brachydanio
34	425	17.0	445	13	Q9W6B3	Q9W6B3 colurix co
35	424	17.0	452	13	Q91155	Q91155 notophthalm
36	424	17.0	459	11	Q9QWJ1	Q9QWJ1 rattus norv
37	422	16.9	416	13	Q9W785	Q9W785 salmo salar
38	422	16.9	448	13	Q9Z019	Q9Z019 xenopus lae
39	421.5	16.9	401	13	Q90382	Q90382 calirina mos
40	418	16.8	410	6	Q97715	Q97715 sus scrofa
41	418	16.8	453	13	Q90967	Q90967 gallus gall
42	416.5	16.7	418	4	Q9P300	Q9P300 homo sapien
43	416	16.7	460	18	Q90966	Q90966 gallus gall
44	416	16.7	462	11	P97513	P97513 mus spretus
45	416	16.7	549	13	Q90970	Q90970 gallus gall

ALIGNMENTS

RESULT 1

ID	Q9UNW4	PRELIMINARY:	PRT:	473 AA.
AC	Q9UNW4			
DT	01-MAY-2000 (TREMBL)	13, Created		
DT	01-MAY-2000 (TREMBL)	13, Last sequence update		
DT	01-OCT-2000 (TREMBL)	15, Last annotation update		
DE	ORPHAN NUCLEAR RECEPTOR.			
GN	PAR2.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=LIVER.			
RX	MEDLINE=8445350; PubMed=9770465;			
RA	Bertilsson G., Heidrich J., Svensson K., Asman M., Jendeberg L.,			
RA	Sydow-Backman M., Ohlsson R., Postlind H., Blomquist P.,			
RA	Berkestan A.;			
RT	"Identification of a human nuclear receptor defines a new signaling			
RT	pathway for CYP3A induction."			
RL	Proc. Natl. Acad. Sci. U.S.A. 95:12208-12213(1998).			
DR	EMBL; AF084644; AAC6457.1; -.			
DR	HSSP; P10826; IHRA.			
DR	INTERPRO; IPR000324; -.			
DR	INTERPRO; IPR000536; -.			
DR	INTERPRO; IPR001628; -.			
DR	INTERPRO; IPR001723; -.			
DR	PFAM; PF00104; hormone_rec; 1.			
DR	PFAM; PF00105; zf-C4; 1.			
DR	PRINTS; PRO0047; STROIDFINGER.			
DR	PRINTS; PRO0350; VITAMINDR.			
DR	PRINTS; PRO0398; STRDHOMONR.			
DR	PROSITE; PS00031; NUCLEAR_RECEPTOR; 1.			
KW	Receptor.			
SO	SEQUENCE			

Query Match: 100.0%; Score 2494; DB 4; Length 473;
Best Local Similarity: 100.0%; Pred. No. 3.7e+204;
Matches 473; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Best Local Similarity 98.4%; Pred. No. 3,2e-190;
Matches 442; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

```

QY 1 MTVTTHHKEGSLRAPALPLHSAALASNPGRPEANLEVRKESWNHADVHCEDTE 60
DB 1 MTVTTHHKEGSLRAPALPLHSAALASNPGRPEANLEVRKESWNHADVHCEDTE 60
QY 61 SVPGKPSVNADEEVGGPOICRVCGKATGYHENVMTCEGCGFFRRAMKRNALRCPEPK 120
DB 61 SVPGKPSVNADEEVGGPOICRVCGKATGYHENVMTCEGCGFFRRAMKRNALRCPEPK 120
QY 121 GACETTRTRROCOARLRCLESGMKKEMTSDAVERRALIKRKKSERTGTQPLGVQ 180
DB 121 GACETTRTRROCOARLRCLESGMKKEMTSDAVERRALIKRKKSERTGTQPLGVQ 180
QY 181 GLTEQRMARIRLMDAQMTFDTTSHFKNFRLPGVLSGCELPESLQAPREBAKMSQ 240
DB 181 GLTEQRMARIRLMDAQMTFDTTSHFKNFRLPGVLSGCELPESLQAPREBAKMSQ 240
QY 241 VKRDLCSLKVSLQNGEDGSVWNYKPPADSGGKEIFSLPHMADSTYMFKGIIISFAKYI 300
DB 241 VKRDLCSLKVSLQNGEDGSVWNYKPPADSGGKEIFSLPHMADSTYMFKGIIISFAKYI 300
QY 301 SFERDLPEDQISLKGAFELCOLRNTVFNAETGTEGCRSLCYLEDTAGGFQQLLE 360
DB 301 SFERDLPEDQISLKGAFELCOLRNTVFNAETGTEGCRSLCYLEDTAGGFQQLLE 360
QY 361 PMLKHYMLKKQLHHEEYVLMQALISLSPDPGVLQHRVVDQLOEOPATILKSYIECNR 420
DB 361 PMLKHYMLKKQLHHEEYVLMQALISLSPDPGVLQHRVVDQLOEOPATILKSYIECNR 420
QY 421 POPARPFLEKIMAMTELRSINAOHTORLRIODIHFPATPLMOELFGITGS 473
DB 421 POPARPFLEKIMAMTELRSINAOHTORLRIODIHFPATPLMOELFGITGS 473

RESULT 2
Q9UJ26 PRELIMINARY; PRT; 457 AA.
AC Q9UJ26;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE NUCLEAR HORMONE RECEPTOR PRRI-C.
GN PRRI.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RA Heard D.J., Holloway J., Hansen C., Tommerup N., Aagaard L.,
RA Vissing H.;
RT "Identification of a novel protein isoform of the human nuclear
RT hormone receptor PXR/SXR and localization to chromosome 3q12.1
RT -13.3."
RL Eur. J. Hum. Genet. 0:0-0(0).
DR EMBL; AJ009936; CAB55490.1;
DR HSSP; P10826; IHRA.
DR INTERPRO; IPR000324;
DR INTERPRO; IPR000536;
DR INTERPRO; IPR001628;
DR INTERPRO; IPR001723;
DR PFAM; PF00104; hormone_rec. 1.
DR PRINTS; PR00047; STEROIDFINGER.
DR PRINTS; PR00350; VITAMINDR.
DR PRINTS; PR00398; STRDHOMONER.
DR PROSITE; PS00031; NUCLEAR_RECEPTOR; 1.
KW Receptor.
SQ SEQUENCE 457 AA; 52145 MW; 812F8B354B5B784E CRC64;

```

Query Match 93.4%; Score 2330; DB 4; Length 457;

Best Local Similarity 98.4%; Pred. No. 3,2e-190;
Matches 442; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

```

QY 25 AAELASHHPGPANLEVRKESWNHADVHCEDTESVPGKPSVNADEEVGGPOICRVCG 84
DB 9 AKMLPPNSPGRPEANLEVRKESWNHADVHCEDTESVPGKPSVNADEEVGGPOICRVCG 68
QY 85 DKATGYFENVMTCEGCGFFRRAMKRNALRCPEFGACETTRTRROCOARLRCLES 144
DB 69 DKATGYFENVMTCEGCGFFRRAMKRNALRCPEFGACETTRTRROCOARLRCLES 129
QY 145 GMRKEMTSDAVERRALIKRKKSERTGTQPLGVQGLTEQRMARIRLMDAQMTFDTT 204
DB 129 GMRKEMTSDAVERRALIKRKKSERTGTQPLGVQGLTEQRMARIRLMDAQMTFDTT 188
QY 205 FSHFKNFRLPGVLSGCELPESLQAPREBAKMSQVRKDLCSLKVSLQNGEDGSVWNY 264
DB 189 FSHFKNFRLPGVLSGCELPESLQAPREBAKMSQVRKDLCSLKVSLQNGEDGSVWNY 248
QY 265 KPPADSGGKEIFSLPHMADSTYMFKGIIISFAKYISFERDLPEDQISLKGAFELCO 324
DB 249 KPPADSGGKEIFSLPHMADSTYMFKGIIISFAKYISFERDLPEDQISLKGAFELCO 308
QY 325 LRNTVFNAETGTEGCRSLCYLEDTAGGFQQLLEPMLKHYMLKKQLHHEEYVLMQ 384
DB 309 LRNTVFNAETGTEGCRSLCYLEDTAGGFQQLLEPMLKHYMLKKQLHHEEYVLMQ 368
QY 385 ISLSPDPGVLQHRVVDQLOEOPATILKSYIECNRPOPANRPFLEKIMAMTELRSINA 444
DB 369 ISLSPDPGVLQHRVVDQLOEOPATILKSYIECNRPOPANRPFLEKIMAMTELRSINA 428
QY 445 QHTORLRIODIHFPATPLMOELFGITGS 473
DB 429 QHTORLRIODIHFPATPLMOELFGITGS 457

```

Best Local Similarity 98.4%; Pred. No. 3,2e-190;
Matches 442; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

```

RESULT 3
Q9UJ27 PRELIMINARY; PRT; 434 AA.
AC Q9UJ27;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE NUCLEAR HORMONE RECEPTOR PRRI-A.
GN PRRI.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RA Heard D.J., Holloway J., Hansen C., Tommerup N., Aagaard L.,
RA Vissing H.;
RT "Identification of a novel protein isoform of the human nuclear
RT hormone receptor PXR/SXR and localization to chromosome 3q12.1
RT -13.3."
RL Eur. J. Hum. Genet. 0:0-0(0).
DR EMBL; AJ009936; CAB55489.1;
DR HSSP; P10826; IHRA.
DR INTERPRO; IPR000324;
DR INTERPRO; IPR000536;
DR INTERPRO; IPR001628;
DR INTERPRO; IPR001723;
DR PFAM; PF00104; hormone_rec. 1.
DR PRINTS; PR00047; STEROIDFINGER.
DR PRINTS; PR00350; VITAMINDR.
DR PRINTS; PR00398; STRDHOMONER.
DR PROSITE; PS00031; NUCLEAR_RECEPTOR; 1.
KW Receptor.
SQ SEQUENCE 434 AA; 49777 MW; AB36DEC23C4C4200 CRC64;

```


DR PRINTS; PRO0047; STROIDFINGER.
 DR PRINTS; PRO0350; VITAMINDR.
 DR PRINTS; PRO0398; STRDHOMONER.
 DR PRINTS; PRO0546; THYROIDHOMR.
 DR PROSITE; PS00031; NUCLEAR_RECEPTOR; 1.
 DR RECEPTOR.
 KW SEQUENCE 397 AA; 45782 MW; 5E2CE0BB10663073 CRC64;

Query Match 82.5%; Score 2058.5; DB 4; Length 397;
 Best Local Similarity 90.8%; Pred. No. 3,5e-167;
 Matches 394; Conservative 1; Mismatches 2; Indels 37; Gaps 1;

QY 40 LEVAPKESWNADYVHEDTESVPGKPSVNADEEVGPOICRYCGDRTATGTHFNVMTCEG 99
 :|||||
 DB 1 MEVAPKESWNNTDFVHCEDESVPKPSVNADEEVGPOICRYCGDRTATGTHFNVMTCEG 60
 QY 100 CKGFRRAMKRNALRCPPFRKACETRTKTRQCCACRLKCKESGKMKMSDEAVEE 159
 :|||||
 DB 61 CKGFRRAMKRNALRCPPFRKACETRTKTRQCCACRLKCKESGKMKMSDEAVEE 120
 QY 160 RRALKRKSRGTOTPLGVQGLTEBQRMIRLMDAOMKTPDTTSHKRNRLPGVLS 219
 :|||||
 DB 121 RRALKRKSRGTOTPLGVQGLTEBQRMIRLMDAOMKTPDTTSHKRNRLPGVLS 173
 QY 220 GCELPESLQAPSRBEAKMSQVRKDCSLKVSQLRGEGDSVWNYKPPADSGGKEIFSL 279
 :|||||
 DB 174 -----VSQLRGEGDSVWNYKPPADSGGKEIFSL 203
 QY 280 PHMADSTYFKGIISPAKYSYFRDLPIDQISLKGAFELCOLRENTVNAETGWE 339
 :|||||
 DB 204 PHMADSTYFKGIISPAKYSYFRDLPIDQISLKGAFELCOLRENTVNAETGWE 263
 QY 340 CGRLSTYCEPAGFQOLLEPMKRFHYMLKQLHBEVYVLMQALSFSPPRGVLOHR 399
 :|||||
 DB 264 CGRLSTYCEPAGFQOLLEPMKRFHYMLKQLHBEVYVLMQALSFSPPRGVLOHR 323
 QY 400 VVDQLOQFOAITLKSYTECNRPQAHRFELFKIMAMLTLSINAGTQRLRIODIRIQ 459
 :|||||
 DB 324 VVDQLOQFOAITLKSYTECNRPQAHRFELFKIMAMLTLSINAGTQRLRIODIRIQ 383
 QY 460 ATPLMOELFGITGS 473
 :|||||
 DB 384 ATPLMOELFGITGS 397

RESULT 6
 ID 09U025 PRELIMINARY; PRT; 379 AA.

AC 09U025; 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
 DE NUCLEAR HORMONE RECEPTOR PRAL-B.
 GN PRAL.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=LIVER;
 RA Heard D.J., Holloway J., Hansen C., Tommerup N., Aagaard L.,
 RA Vissing H.;
 RT "Identification of a novel protein isoform of the human nuclear
 RT hormone receptor PXR/SXR and localization to chromosome 3q12.1
 RT -13.3.";
 RL Eur. J. Hum. Genet. 0:0-0(0).
 DR EMBL; AF009936; CAB55491.1;
 DR HSSP; P10826; 1HRA.
 DR INTERPRO; IPR000324;
 DR INTERPRO; IPR000536;
 DR INTERPRO; IPR001628; -;

DR INTERPRO; IPR001723; -;
 DR PFAM; PF00104; hormone_rec. 1.
 DR PFAM; PF00105; zf-CA. 1.
 DR PRINTS; PRO0047; STROIDFINGER.
 DR PRINTS; PRO0350; VITAMINDR.
 DR PRINTS; PRO0398; STRDHOMONER.
 KW RECEPTOR.
 KW SEQUENCE 379 AA; 43692 MW; 8F7B7AFA13E45036 CRC64;

Query Match 79.1%; Score 1973; DB 4; Length 379;
 Best Local Similarity 99.7%; Pred. No. 6.2e-160;
 Matches 378; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 95 MTCEGCKGFRRAMKRNALRCPPFRKACETRTKTRQCCACRLKCKESGKMKMSD 154
 :|||||
 DB 1 MTCEGCKGFRRAMKRNALRCPPFRKACETRTKTRQCCACRLKCKESGKMKMSD 60
 QY 155 EAVEERALLIKRKSERTOTPLGVQGLTEBQRMIRLMDAOMKTPDTTSHKRNRLP 214
 :|||||
 DB 61 EAVEERALLIKRKSERTOTPLGVQGLTEBQRMIRLMDAOMKTPDTTSHKRNRLP 120
 QY 215 GYLSGCELPESLQAPSRBEAKMSQVRKDCSLKVSQLRGEGDSVWNYKPPADSGGKE 274
 :|||||
 DB 121 GYLSGCELPESLQAPSRBEAKMSQVRKDCSLKVSQLRGEGDSVWNYKPPADSGGKE 180
 QY 275 IFSLPHMADSTYFKGIISPAKYSYFRDLPIDQISLKGAFELCOLRENTVNAE 334
 :|||||
 DB 181 IFSLPHMADSTYFKGIISPAKYSYFRDLPIDQISLKGAFELCOLRENTVNAE 240
 QY 335 TGTEWCGRLSTYCEPAGFQOLLEPMKRFHYMLKQLHBEVYVLMQALSFSPPRG 394
 :|||||
 DB 241 TGTEWCGRLSTYCEPAGFQOLLEPMKRFHYMLKQLHBEVYVLMQALSFSPPRG 300
 QY 395 VLOHNVVDQLOQFOAITLKSYTECNRPQAHRFELFKIMAMLTLSINAGTQRLRIQ 454
 :|||||
 DB 301 VLOHNVVDQLOQFOAITLKSYTECNRPQAHRFELFKIMAMLTLSINAGTQRLRIQ 360
 QY 455 DIHPATPLMOELFGITGS 473
 :|||||
 DB 361 DIHPATPLMOELFGITGS 379

RESULT 7
 ID 09T002 PRELIMINARY; PRT; 411 AA.

AC 09T002; 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
 DE PREGNANE X RECEPTOR.
 GN NR112 OR PXR.
 OS Oryctolagus cuniculus (Rabbit).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 OX NCBI_TaxID=9986;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRED NEW ZEALAND WHITE RABBIT; TISSUE=KIDNEY;
 RA Savas U., Wester M.R., Griffin K.J., Johnson E.F.;
 RT "The rabbit pregnane x receptor is activated by rifampicin";
 RT submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
 RL [2]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=20092326; PubMed=10628745;
 RA Jones S.A., Moore L.B., Shenk J.L., Wisely G.B., Hamilton G.A.,
 RA McKee D.D., Tomkinson N.C., LeCluyse E.L., Lambert M.H., Willson T.M.,
 RA Klierer S.A., Moore J.T.;
 RT "The pregnane x receptor: a promiscuous xenobiotic receptor that has
 RT diverged during evolution";
 RL Mol. Endocrinol. 14:27-39(2000).
 DR EMBL; AF182217; AAD54426.1;
 DR EMBL; AF188476; AAF31165.1; -;

DR INTERPRO: IPR001628; -
 DR INTERPRO: IPR001723; -
 DR INTERPRO: IPR001728; -
 DR PFAM: PF00104; hormone_rec; 1.
 DR PFAM: PF00105; zf-C4; 1.
 DR PRINTS: PR00047; STEROIDFINGER.
 DR PRINTS: PR00350; VITAMINDR.
 DR PRINTS: PR00398; STRDHOMOMER.
 DR PRINTS: PR00546; THYROIDHOMR.
 DR RECEPTOR.
 KM SEQUENCE 342 AA; 39696 MW; 7C9467C3E9000A3A CRC64;

Query Match 70.3%; Score 1753.5; DB 4; Length 342;
 Best Local Similarity 90.0%; Pred. No. 2.6e-141;
 Matches 341; Conservative 0; Mismatches 1; Indels 37; Gaps 1;

QY 95 MTCEGCKGFFRRAMRNRLRCFPRKACETIRKTRROCOACRLKCEGSKMKEMMSD 154
 DB 1 MTCEGCKGFFRRAMRNRLRCFPRKACETIRKTRROCOACRLKCEGSKMKEMMSD 60
 QY 155 EAVEERRALIKRKSEKERTGTPGLGVQGLTEEQRMIRRLMDAQMTFTTSHKNEFLP 214
 DB 61 EAVEERRALIKRKSEKERTGTPGLGVQGLTEEQRMIRRLMDAQMTFTTSHKNEFLP 118
 QY 215 GVLSSGCELPESLQAPSRERAKNSQVRKDCSLKVSQINGEGOSVWNYKPPADSGGKE 274
 DB 119 -----VSLQLNGEDSSVWNYKPPADSGGKE 143
 QY 275 IFSLLPHMADMTYMEKGLISPAKISYFRDLPIEDQISLKGAFELCOLRENTVNAE 334
 DB 144 IFSLLPHMADMTYMEKGLISPAKISYFRDLPIEDQISLKGAFELCOLRENTVNAE 203
 QY 335 TGTWEGRLSYCLEDTAGGPOQLLEPMKRYMMLKQLHREEVYLMQALISLSPDRG 394
 DB 204 TGTWEGRLSYCLEDTAGGPOQLLEPMKRYMMLKQLHREEVYLMQALISLSPDRG 263
 QY 395 VLOHNVQLOBOFAITLKSYTECNRPQAHFELFKIMAMTELRSINAQHTQRLRIQ 454
 DB 264 VLOHNVQLOBOFAITLKSYTECNRPQAHFELFKIMAMTELRSINAQHTQRLRIQ 323
 QY 455 DIHPFATPLMOELFGITGS 473
 DB 324 DIHPFATPLMOELFGITGS 342

RESULT 10
 Q91839 PRELIMINARY: PRT: 386 AA.
 AC Q91839.
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
 DE ORPHAN NUCLEAR RECEPTOR OF STEROID/THYROID SUPERFAMILY.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
 OC Xenopodinae; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94173664; PubMed=8127656;
 RA Smith D.P., Mason C.S., Jones E.A., Old R.W.;
 RT "A novel nuclear receptor superfamily member in Xenopus that
 RT associates with RXR, and shares extensive sequence similarity to the
 RT mammalian vitamin D3 receptor.";
 RL Nucleic Acids Res. 22:66-71(1994).
 DR EMBL: X75163; CA53006.1; -
 DR HSP: P10826; IIRA.
 DR INTERPRO: IPR000536; -
 DR INTERPRO: IPR001628; -
 DR PFAM: PF00104; hormone_rec; 1.
 DR PFAM: PF00105; zf-C4; 1.

DR PRINTS: PR00047; STEROIDFINGER.
 DR PROSITE: PS00031; NUCLEAR RECEPTOR: UNKNOWN 1.
 SQ SEQUENCE 386 AA; 44302 MW; 8B046ED219556573 CRC64;

Query Match 39.8%; Score 991.5; DB 13; Length 386;
 Best Local Similarity 49.0%; Pred. No. 2.3e-76;
 Matches 204; Conservative 59; Mismatches 104; Indels 49; Gaps 9;

QY 57 EDTESVPGKPSVNADEEVGPOICRVGDKATGYHNVMTGCGCGFFRRAMRNRLRC 116
 DB 14 EEEEDASNSCGTEGDEDDDPKICRACGDRATGYHNVMTGCGCGFFRRAMRNRLRC 73
 QY 117 PFRKACETIRKTRROCOACRLKCEGSKMKEMMSDAVEERRALIKRK-KSERGTGQ 175
 DB 74 PFRKACETIRKTRROCOACRLKCEGSKMKEMMSDAVEERRALIKRK-KSERGTGQ 132
 QY 176 PLGVQGLTEEQRMIRRLMDAQMTFTTSHKNEFLPGVLSGCELPESLQAPSRERA 235
 DB 133 PLGVQGLTEEQRMIRRLMDAQMTFTTSHKNEFLPGVLSGCELPESLQAPSRERA 171
 QY 236 AKMSQVRKDCSLKVSQINGEGOSVWNYKPPADSGGKEIFSLLPHMADMTYMEKGLIS 295
 DB 172 -----RSSDPT--OEPOATS--SEAFMLPLHISDLVTYMKGLIS 207
 QY 296 FAVISYFRDLPIEDQISLKGAFELCOLRENTVNAETGTWEGRLSYCLEDT-AGGE 354
 DB 208 FAVISYFRDLPIEDQISLKGAFELCOLRENTVNAETGTWEGRLSYCLEDT-AGGE 267
 QY 355 QQLLEPMKRYMMLKQLHREEVYLMQALISLSPDRGVLOHNVQLOBOFAITLKS 414
 DB 268 QQLLEPMKRYMMLKQLHREEVYLMQALISLSPDRGVLOHNVQLOBOFAITLKS 327
 QY 415 YICNRPQAHFELFKIMAMTELRSINAQHTQRLRIQDIHPFATPLMOELFG 469
 DB 328 YICNRPQAHFELFKIMAMTELRSINAQHTQRLRIQDIHPFATPLMOELFG 383

RESULT 11
 Q91873 PRELIMINARY: PRT: 420 AA.
 AC Q91873.
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
 DE VITAMIN D RECEPTOR A.
 OS Paralicthys olivaceus (Flounder).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorpha; Acanthopterygii; Percomorpha; Pleuronectiformes;
 OC Pleuronectoidae; Bothidae; Paralicthys.
 OX NCBI_TaxID=8255;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=INTESTINE;
 RA Suzuki T., Suzuki N., Srivastava A.S., Kurokawa T.;
 RT "Identification of cDNAs encoding two subtypes of vitamin D receptor
 RT in flounder, Paralicthys olivaceus.";
 RL Biochem. Biophys. Res. Commun. 270:40-45(2000).
 DR EMBL: AB037674; BAA95016.1; -
 KM RECEPTOR.
 SQ SEQUENCE 420 AA; 47486 MW; 038FBF00D4F38067 CRC64;

Query Match 32.8%; Score 817; DB 13; Length 420;
 Best Local Similarity 43.1%; Pred. No. 1.8e-61;
 Matches 179; Conservative 70; Mismatches 144; Indels 22; Gaps 12;
 QY 68 VNADE-EVGGPOICRVGDKATGYHNVMTGCGCGFFRRAMRNRLRCFPRKACETIT 126
 DB 11 VGPDEFRRNRLRICGVGDKATGYHNVMTGCGCGFFRRAMRNRLRCFPRKACETIT 69

```

QY 127 RKTROCCACLRKCSGKMKKEMINSDEAVEERRALLIKRKKSERTGTGLOPLAGVGTLEBO 186
DB 70 KDNRRHCACLRKCRIDIGMKFEILTDEVOKEKELIKRKEEAREAMPR-LNEQO 128
QY 187 RMIRBELMDQMKTEQDTFSHFKEFLP---GYLSSGCELPESLOAPSPBEAARNSQVRK 243
DB 129 ARMTSSLEAHMKKYDASYSDFSRFPVREGVTRASRAASLHSLSDASSDSFVHSPE 188
QY 244 DLSCLAV---SLDRGEDGSVWYKPPADSGKEIFSLPHMAADSTYFKGIIISPAVI 300
DB 189 SV-DTKNFSLLMYODGA-----SPDSSEETKLSMLPHLADLVYSIQVYIGFAKMT 243
QY 301 SYFDLPEDIOISLKGAELCOLARNTVFNAGTWCQ--RLSYCLD--TAGGFOOL 357
DB 244 PGFRLDLAEDQIALKSSAIEIIMRSNOSFLEDSWSCGGDPFKCINDVTKAGHILE 303
QY 358 LLEPMLKFMHMKLQHEEYVLMQALISFSPDRGVQLQHRVQDLOQFATTKSIE 417
DB 304 LLEPLVAFQVGLKRLNHEEHVLMGICLSPDRGVQDHAHVQDRLPEALQAVIR 363
QY 418 CNRPOAHRELFLKIMAMLTSLRSINAQHTQ--RLLRIDIRPF-ATPLMOELFG 469
DB 364 IN--HFGGRLLVAKMIQKLDLRLSNEHSHKQYRSLSPQHSKQUTPLVLEVFG 416

```

```

RESULT 12
Q9PTN2 PRELIMINARY; PRT: 453 AA.
AC 09PTN2;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE VITAMIN D RECEPTOR.
GN VDR.
OS Brachydanio rerio (zebrafish) (zebra danio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
OC Cypriniformes; Cyprinidae; Rasbora; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RA Koumenko A.P.;
RT Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF164512; AAF21427.1;
DR HSSP; P20393; IABY.
DR INTERPRO; IPR000324;
DR INTERPRO; IPR000536;
DR INTERPRO; IPR001628;
DR INTERPRO; IPR001723;
DR PFAM; PF00104; hormone_rec; 1.
DR PFAM; PF00105; zf-C4; 1.
DR PRINTS; PR00047; STROIDFINGER.
DR PRINTS; PR00350; VITAMINDR.
DR PRINTS; PR00398; STRDHOMOMER.
DR PROSITE; PS00031; NUCLEAR_RECEPTOR; 1.
KW Receptor.
SQ SEQUENCE 453 AA: 50765 MW; B52C92C185859F9C CRC64;

```

```

Query Match 31.9%; Score 796; DB 13; Length 453;
Best Local Similarity 39.7%; Pred. No. 1.2e-59;
Matches 180; Conservative 76; Mismatches 165; Indels 32; Gaps 11;
QY 35 GEPANLEVPKESWNAHDFVHCEDTESVGRPSYNADEVGPGQICRVGDKATGYFNV 94
DB 11 GKSCEAGACSESVNGDATSLMDLMAVSTATGDDPDRNAPICGCGDKATGFHFA 70
QY 95 MTCSECKGCFERAMRNRLCSPFKAGCETTKTRROCAQCLRKLESQKMKEMTSD 154
DB 71 MTCSECKGCFERAMRNRLCSPFKAGCETTKTRROCAQCLRKLESQKMKEMTSD 154
QY 155 EAVEERRALLIKRKKSE--RTGTGLOPLAGVGTLEBO RMIRBELMDQMKTEQDTFSHFKNF 211

```

```

DB 130 EYQRRKDLIMKKEEAREAKRP---RLSDEQMLTINSVYEAHNTYDSDYFVR 185
QY 212 RLP---GYLSSGCELPESLOAPSPBEAARNS-----QVRKDLCLSKSLDRGEDGSVW 262
DB 186 RPPVREGPVTRASRAASLHSLSDASSDSFNSPESVDTKLNFSLMLMYODSGSPDS-- 243
QY 263 NYRPPADSGKEIFSLPHMAADSTYFKGIIISPAKVISYFDLPEDIOISLKGAEL 322
DB 244 ----SEEDQNSRLSMLPHLADLVYSIQVYIGFAKMTPGFRLDLAEDQIALKSSAIEI 298
QY 323 COLRFTVFNAGTWCQ--RLSYCLD--TAGGFOOLLEPMLKFMHMKLQHEEY 379
DB 299 IMRSNOSFLEDSWSCGGDPFKCINDVTKAGHILELLEPLVAFQVGLKRLNHEEH 358
QY 380 VLMQALISFSPDRGVQLQHRVQDLOQFATTKSIECNRPOAHRELFLKIMAMLTSL 439
DB 359 VILMAICLSPDRGVQDHAHVQDRLPEALQAVIRIQ--HFGGRLLVAKMIQKLDL 416
QY 440 RSINAQHTQ--RLLRIDIRPF-ATPLMOELFG 469
DB 417 RSLNEHSHKQYRSLSPQHSKQUTPLVLEVFG 449

```

```

RESULT 13
Q9IB74 PRELIMINARY; PRT: 425 AA.
AC 09IB74;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE VITAMIN D RECEPTOR B.
GN VDRB.
OS Parachanna olivacea (Flounder).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthopterygia; Acanthopterygii; Perciformes; Pleuronectiformes;
OC Pleuronectidae; Bothidae; Parachanna.
OX NCBI_TaxID=8255;
RN [1]
RP SEQUENCE FROM N.A.
RA TISSUE-INTESTINE;
RA Suzuki T., Suzuki N., Sivasubava A.S., Kurokawa T.;
RT Identification of cDNAs encoding two subtypes of vitamin D receptor
RT in flounder, Parachanna olivacea.
DR Biochem. Biophys. Res. Commun. 270:40-45(2000).
DR EMBL; AB037673; BAA95015.1;
KW Receptor.
SQ SEQUENCE 425 AA: 48174 MW; 7B1D49CF4C2C65E9 CRC64;

```

```

Query Match 31.7%; Score 790.5; DB 13; Length 425;
Best Local Similarity 41.6%; Pred. No. 3.3e-59;
Matches 179; Conservative 78; Mismatches 140; Indels 33; Gaps 13;
QY 61 SVPGPSYNADE-ENVGPGQICRVGDKATGYFNVMTCECKGCFERAMRNRLCSPR 119
DB 4 TVYTSLSASDEPDRNMPRICGCGDKATGFHFAHMAKCECKGCFERAMRNRLCSPR 62
QY 120 KGACETTRKTRROCAQCLRKLESQKMKEMINSDEAVEERRALLIKRKK-----SERT 172
DB 63 NQSCITTDNRHRCACLRKCRIDIGMKFEILTDEVOKEKELIKRKEEAREAMPR-LNEQO 128
QY 173 GTQPLGVGLTEBO RMIRBELMDQMKTEQDTFSHFKEFLP---GYLSSGCELPESLOA 229
DB 123 ARP---RLSDEQNSRLSMLPHLADLVYSIQVYIGFAKMTPGFRLDLAEDQIALKSSAIEI 298
QY 230 PSREBAARNSQVRKDLCLSKV-----LQLRGEDGSVWYKPPADSGKEIFSLPHMAAD 285
DB 179 LSDASSDSFNSPESV-DTKNFSLLMYODG-----SPDSSEETKLSMLPHLADLV 243
QY 286 STYFKGIIISPAKVISYFDLPEDIOISLKGAELCOLARNTVFNAGTWCQ--L 343

```

```

Db 234 VESIOKIVIGFAMKIPGRELTAEDQIALKSSAEVIMLNSQNFJEDMSGCGADP 293
QY 344 SYCLED-TAGGFOQLLEPMLEKFKHMKLOLHEEYVLMQALISPSDRPGVLOHRVD 402
Db 294 KJOISVYTAGHTLELLELVFQVGLKTLNLOEBEHVLMALICLSPDRPGVQCHARLE 353
QY 403 OLOEPAITLKSYIECNRPDPARHLEFLKIMAMTELKSNQHO--RLRIDIHFF- 459
Db 354 ALQDRLESETLQAVIQH--HPGRLLYAMKIQKLADRLSLNEHSHKQYRSLSFRESHQ 411
QY 460 APPLMEELFG 469
Db 412 LPPLELVSSG 421

RESULT 14
ID 090US1 PRELIMINARY; PRT: 358 AA.
AC 090US1;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE NUCLEAR RECEPTOR.
GN CAR.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SOURCE FROM N.A.
RC STRAIN=WISTAR KYOTO, AND FISCHER;
RA Yoshinari K., Sueyoshi T., Moore R., Negishi M.;
RT "Sexually dimorphic nuclear translocation of receptor CAR and
RT induction of CYP2B1 gene by phenobarbital in rat livers.";
RT Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
RL EMBL: A133095; AAF22567.1; -
DR EMBL: A133094; AAF22566.1; -
DR HSSP: P19793; 2N1L.
DR INTERPRO: IPR000324; -
DR INTERPRO: IPR000536; -
DR INTERPRO: IPR001628; -
DR INTERPRO: IPR001723; -
DR INTERPRO: IPR001728; -
DR PFAM: PF00104; hormone_rec. 1.
DR PFAM: PF00105; zf-c4; 1.
DR PRINTS: PR00047; STROIDFINGER.
DR PRINTS: PR00350; VITAMINDR.
DR PRINTS: PR00398; STRDHOMOMR.
DR PRINTS: PR00546; THYROIDHOMR.
DR PROSITE: PS00031; NUCLEAR_RECEPTOR; 1.
KW Receptor.
SQ SEQUENCE 358 AA; 40922 MW; 13691F49CAD8F1ED CRC64;

Query Match 29.2%; Score 729; DB 11; Length 358;
Best Local Similarity 40.3%; Pred. No. 4.5e-54;
Matches 160; Conservative 61; Mismatches 122; Indels 54; Gaps 6;

```

```

QY 312 ISILKGAFFCOLRENTVENAFETWEGRLSYCLEDIAG-GFOQLLEPMLEKFKHMK 370
Db 201 ISILKGAFFILHISLNTTTCLOTQNFPGCPICRYKEDAVNHGFOYELELIHFKTLK 260
QY 371 KLOLHEEYVLMQALISPSDRPGVLOHRVDLOBOFAITLKSYIECNRPDPARHLEFL 430
Db 261 RLQOEPEYALMAMALFSPDRPGVTOREIDLOEVALILNNHIMEQOSRLQSRFLYA 320
QY 431 KIMAMTELKSNQHOQRLRTQDIHFFATPLMOEL 467
Db 321 KIMGLAELKSNQSYSEIHRIOGLSAM-MPLUGEL 356

RESULT 15
ID 090UH7 PRELIMINARY; PRT: 445 AA.
AC 090UH7;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE NUCLEAR OXYSTEROL RECEPTOR LXR-ALPHA (LXR-ALPHA).
GN NR1H3 OR LXR-ALPHA.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SOURCE FROM N.A.
RC STRAIN=129SV; TISSUE=LIVER;
RA Albert S., Steffensen K.R., Gustafsson J.A.;
RT "Cloning and characterization of nuclear oxysterol receptor genes LXR
RT and LXRb from mouse.";
RT Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
RL EMBL: A132599; CAB51952.1; -
DR EMBL: A132600; CAB51952.1; JOINED.
DR EMBL: A132601; CAB51923.1; -
DR HSSP: P20393; 1A6Y.
DR MGD: MGI:1352462; Nr1h3.
DR INTERPRO: IPR000324; -
DR INTERPRO: IPR000536; -
DR INTERPRO: IPR000923; -
DR INTERPRO: IPR001628; -
DR INTERPRO: IPR001723; -
DR INTERPRO: IPR001728; -
DR INTERPRO: IPR003069; -
DR INTERPRO: IPR003078; -
DR PFAM: PF00104; hormone_rec. 1.
DR PFAM: PF00105; zf-c4; 1.
DR PRINTS: PR00047; STROIDFINGER.
DR PRINTS: PR00350; VITAMINDR.
DR PRINTS: PR00398; STRDHOMOMR.
DR PRINTS: PR00546; THYROIDHOMR.
DR PRINTS: PR01283; ECDYSTEROIDR.
DR PRINTS: PR01282; RETINOICACIDR.
DR PROSITE: PS00031; NUCLEAR_RECEPTOR; 1.
DR PROSITE: PS00196; COPPER_BLUE; UNKNOWN_1.
KW Receptor.
SQ SEQUENCE 445 AA; 50476 MW; CC9A8DF38D935593 CRC64;

Query Match 20.2%; Score 504.5; DB 11; Length 445;
Best Local Similarity 30.4%; Pred. No. 7.8e-35;
Matches 130; Conservative 85; Mismatches 142; Indels 71; Gaps 13;

```

QY 171 ---RGTGPTLVGUGITEGORMMIRELMDAOMKTFDTTFSHEKNEFRLPGVLSGCCLEPESL 227
DB 190 VSPRVSSPPQVLPOLSPEDLGMIERKLVAAQCCNRRSES----- 228
QY 228 QAPSRBEAKMSQVRKDLCSLKVSLQLRGEGDSVWNTKPPADSGGKEIFSLPHADNST 287
DB 229 ---DRLRVTFW-PIAPD-----PQSRARQORFA---HFTETAI 260
QY 288 YMFKGIIISPAKVISYFRDLPIDQISLKGAFELCOLFNTVEN--AETGTWEGRLSY 345
DB 261 VSVGEIVDEPAKQLPGLQLSREDQIALKTSALIEVMLETSRRVNPSESITF--LKDSY 319
QY 346 CLEDTA-GGFOQLLEPMLEKHYMLKLQHEEYVLMQAISLESPPDRPGVLQHRVVDL 404
DB 320 NREDEPAKAGIQVEFINPIFEFSRANNEQLNDAEFALLIAISIFSADRPVODQLOVERL 379
QY 405 OEQFAITLKSYIECNRQPAHREFLKIMAMTELRSINAOHTORL--LRIDIHFEATP 462
DB 380 QHTYVEALHAYVINHPH--DRLMFPMLMKLVSLRTLSVHSEGVFALRIQD--KKLP 435
QY 463 LMOELFGI 470
DB 436 LISEIMDV 443

Search completed: February 18, 2001, 13:05:05
Job time: 6194 sec

1	2287	91.7	434	1	PXR_HUMAN	0754619	homo sapien
2	1779.5	71.4	431	1	PXR_MOUSE	054915	mus musculus
3	1761.5	70.6	431	1	PXR_RAT	091187	rattus norv
4	815.5	32.7	448	1	VDR_COTJA	P49701	coturnix co
5	815	32.7	451	1	VDR_CHICK	Q42392	gallus gall
6	814	32.6	422	1	VDR_XENLA	Q13124	xenopus lae
7	801	32.1	423	1	VDR_RAT	P13053	rattus norv
8	792.5	31.8	422	1	VDR_MOUSE	P48281	mus musculu
9	789	31.6	424	1	VDR_BOVIN	028037	bos taurus
10	789	31.6	427	1	VDR_HUMAN	P11473	homo sapien
11	768	30.8	348	1	NR13_HUMAN	Q14994	homo sapien
12	729	29.2	358	1	NR13_RAT	Q95911	rattus norv
13	725	29.1	358	1	NR13_MOUSE	Q35657	mus musculu
14	500.5	20.1	461	1	NRH2_HUMAN	P55055	homo sapien
15	498.5	20.0	445	1	NRH3_MOUSE	Q920V9	mus musculu
16	496.5	19.9	445	1	NRH3_RAT	062685	rattus norv
17	491.5	19.7	447	1	NRH3_HUMAN	Q13133	homo sapien
18	489	19.6	446	1	NRH2_MOUSE	Q60684	mus musculu
19	488	19.6	446	1	NRH2_RAT	062755	rattus norv
20	457.5	18.3	757	1	ECR_LOCCU	Q18531	lucilla cup
21	455.5	18.3	395	1	THB_PAROL	Q91279	paralichthy
22	455	18.2	675	1	ECR_AEPAE	P49880	redes aegypt
23	452.5	18.1	461	1	THB1_HUMAN	Q10828	homo sapien
24	450	18.0	359	1	THB_CHICK	P18112	gallus galli
25	449	18.0	373	1	THB_RANCA	Q02955	rana catesbe
26	448	18.0	373	1	THB_XENLA	P18117	xenopus lae
27	445	17.8	476	1	THB2_HUMAN	P37242	homo sapien
28	444.5	17.8	461	1	THB1_MOUSE	P37242	mus musculu
29	443	17.8	414	1	THBB_XENLA	P18119	xenopus lae
30	443	17.8	461	1	THB1_RAT	P18113	rattus norv
31	442	17.7	475	1	THB2_MOUSE	P37244	mus musculu
32	439.5	17.6	579	1	NRD2_HUMAN	Q14995	homo sapien
33	437	17.5	514	1	THB2_RAT	P37826	rattus norv

DR INTERRO; JFV000330; -

DR INTERPRO: IPR001628; -
 DR INTERPRO: IPR001723; -
 DR PFAM: PF00104; hormone_rec. 1.
 DR PFAM: PF00105; zf-C4; 1.
 DR PRINTS: PR00047; STEROIDFINGER.
 DR PRINTS: PR00350; VITAMINDR.
 DR PRINTS: PR00398; STDRHORMONER.
 DR PROSITE: PS00031; NUCLEAR_RECEPTOR; 1.
 DR Receptor; Transcription regulation; DNA-binding; Nuclear protein;
 Zinc-finger.
 FT DNA_BIND 41 107 C4-TYPE ZINC FINGERS (TWO).
 FT ZN_FING 41 107 C4-TYPE.
 FT ZN_FING 77 62 C4-TYPE.
 FT ZN_FING 108 204 C4-TYPE.
 FT DOMAIN 205 434 HINGE.
 FT SEQUENCE 434 AA; 49761 MW; 1DF6A2AE3109CADA CRC64;

Query Match 91.7%; Score 2287; DB 1; Length 434;
 Best Local Similarity 99.8%; Pred. No. 6,9e-176;
 Matches 433; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 40 LEVPRKSNHADVHCEDESVGKPSVNADEVGGPOICRGCDKATGYHNVMTCEG 99
 DB 1 MEVPRKSNHADVHCEDESVGKPSVNADEVGGPOICRGCDKATGYHNVMTCEG 60
 QY 100 CKGFERRAMKRNALRCPFRKACETTRKTRROCCARLKLKESGKKKEMSDAEVEE 159
 DB 61 CKGFERRAMKRNALRCPFRKACETTRKTRROCCARLKLKESGKKKEMSDAEVEE 120
 QY 160 RRAIKRKKSRGTGTPGVGLTEEDPRMIRLMDAQMFTDTTFSHFNRLPGVLS 219
 DB 121 RRAIKRKKSRGTGTPGVGLTEEDPRMIRLMDAQMFTDTTFSHFNRLPGVLS 180
 QY 220 GCELPESLOAPSRKAKNSOVKRDLSLVSLQREDSVWNYKRPASGKEITFSL 279
 DB 181 GCELPESLOAPSRKAKNSOVKRDLSLVSLQREDSVWNYKRPASGKEITFSL 240
 QY 280 PHMADSTYMEKGIISFAKYSYFRDLPIEDQISLKGAFELCOLRNTVNAETGWE 339
 DB 241 PHMADSTYMEKGIISFAKYSYFRDLPIEDQISLKGAFELCOLRNTVNAETGWE 300
 QY 340 CGRLSTCLEDTAGGFOOLLLEPMKFKHYMKKIQLEHEEYVLMQASISLSPDRPGVLOHR 399
 DB 301 CGRLSTCLEDTAGGFOOLLLEPMKFKHYMKKIQLEHEEYVLMQASISLSPDRPGVLOHR 360
 QY 400 VVDQLOPALTLSYTECHRPQAHRELFLKIMAMTELRSINAQOTGLRIQIDHPR 459
 DB 361 VVDQLOPALTLSYTECHRPQAHRELFLKIMAMTELRSINAQOTGLRIQIDHPR 420
 QY 460 ATPLMQELFGITGS 473
 DB 421 ATPLMQELFGITGS 434

RESULT 2
 PXR_MOUSE STANDARD; PRT; 431 AA.
 ID PXR_MOUSE
 AC 054915;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 01-OCT-2000 (Rel. 40, Last sequence update)
 DE ORPHAN NUCLEAR RECEPTOR PXR (PREGNANE X RECEPTOR).
 GN NR112 OR PXR.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 RN [1]
 RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
 RC TISSUE=LIVER;
 RX MEDLINE=98149345; PubMed=9489701;
 RA Klierer S.A., Moore J.T., Wade L., Staudinger J.L., Watson M.A.,
 Jones S.A., McKee D.D., Oliver B.B., Willson T.M., Zetterstrom R.H.,

RA Perlmann T., Lehmann J.M.;
 RT "An orphan nuclear receptor activated by pregnanes defines a novel
 RT steroid signaling pathway.";
 RL Cell 92:73-82(1998)
 CC -1- FUNCTION: ORPHAN RECEPTOR. ITS NATURAL LIGAND IS PROBABLY
 CC PREGNANE. BINDS TO A RESPONSE ELEMENT IN CYP3A GENES PROMOTER.
 CC -1- SUBUNIT: FORMS A HETERODIMER WITH RXR.
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; 1/PXR.1 (SHOWN HERE) AND
 CC 2/PXR.2; ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- INDUCTION: ACTIVATED BY NATURALLY OCCURRING STEROIDS SUCH AS
 CC PREGNENOLONE AND PROGESTERONE, SYNTHETIC GLUCOCORTICOID AND
 CC ANTIGLUCOCORTICOID AND 16-ALPHA-CARBONITRILE (PCN).
 CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.
 CC NR1 SUBFAMILY.

 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

DR EMBL: AF031814; AAC39964.1; -
 DR MGD: MGI:1337040; NR112.
 DR INTERPRO: IPR000324; -
 DR INTERPRO: IPR000536; -
 DR INTERPRO: IPR001628; -
 DR INTERPRO: IPR001723; -
 DR PFAM: PF00104; hormone_rec. 1.
 DR PFAM: PF00105; zf-C4; 1.
 DR PRINTS: PR00047; STEROIDFINGER.
 DR PRINTS: PR00350; VITAMINDR.
 DR PRINTS: PR00398; STDRHORMONER.
 DR PROSITE: PS00031; NUCLEAR_RECEPTOR; 1.
 DR Receptor; Transcription regulation; DNA-binding; Nuclear protein;
 Zinc-finger; Alternative splicing.
 FT DNA_BIND 38 104 C4-TYPE ZINC FINGERS (TWO).
 FT ZN_FING 38 104 C4-TYPE.
 FT ZN_FING 74 99 C4-TYPE.
 FT ZN_FING 105 201 C4-TYPE.
 FT DOMAIN 202 431 HINGE.
 FT VARSPLIC 171 211 MISSING (IN ISOFORM 2).
 SO SEQUENCE 431 AA; 49567 MW; F592AF91F689329E CRC64;

Query Match 71.4%; Score 1779.5; DB 1; Length 431;
 Best Local Similarity 77.2%; Pred. No. 3.1e-135;
 Matches 332; Conservative 42; Mismatches 55; Indels 1; Gaps 1;

QY 42 VPRKSNHADVHCEDESVGKPSVNADEVGGPOICRGCDKATGYHNVMTCEGCK 101
 DB 1 MPRKSNHADVHCEDESVGKPSVNADEVGGPOICRGCDKATGYHNVMTCEGCK 59
 QY 102 GFFRRAMKRNALRCPFRKACETTRKTRROCCARLKLKESGKKKEMSDAEVEER 161
 DB 60 GFFRRAMKRNALRCPFRKACETTRKTRROCCARLKLKESGKKKEMSDAEVEER 119
 QY 162 ALIKRKKSRGTGTPGVGLTEEDPRMIRLMDAQMFTDTTFSHFNRLPGVLSGGC 221
 DB 120 ALIKRKKSRGTGTPGVGLTEEDPRMIRLMDAQMFTDTTFSHFNRLPGVLSGGC 179
 QY 222 ELPELSQAPSRKAKNSOVKRDLSLVSLQREDSVWNYKRPASGKEITFSLPH 281
 DB 180 ELPELSQAPSRKAKNSOVKRDLSLVSLQREDSVWNYKRPASGKEITFSLPH 239
 QY 282 MADSTYMEKGIISFAKYSYFRDLPIEDQISLKGAFELCOLRNTVNAETGWECC 341
 DB 240 MADSTYMEKGIISFAKYSYFRDLPIEDQISLKGAFELCOLRNTVNAETGWECC 299
 QY 342 RLSTCLEDTAGGFOOLLLEPMKFKHYMKKIQLEHEEYVLMQASISLSPDRPGVLOHRVY 401


```

Query Match          70.6%  Score 1761.5; DB 1; Length 431;
Best Local Similarity 76.7%  Pred. 8.4e-134;
Matches 330; Conservative 42; Mismatches 57; Indels 1; Gaps
0Y  42 VRPKSEWNHADVHCEDFESYPGKPSVYNNDEVGAPOTCRVCGDKATGYHFFVYMTCEGCK 10Y
:|||||:||||| 1 1: :||| 1 : : ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

DN      01-FEB-1996 (Rel. 33, Created)
DT      01-FEB-1996 (Rel. 33, Last sequence update)
DT      30-MAY-2000 (Rel. 39, Last annotation update)
DE      VITAMIN D3 RECEPTOR (VDR) (1,25-DIHYDROXYVITAMIN D3 RECEPTOR).
GN      VDR OR NR1J1.
OS      Coturnix coturnix japonica (Japanese quail).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC      Coturnix.
RN      [1]
RP      SEQUENCE FROM N.A.
RC      TISSUE=CHORIOALLANTOIC MEMBRANE;
RX      MEDLINE=95062315; PubMed=7972109;
RA      Elaroussi M.A., Prah J.M., Deluca H.F.;
RT      "The avian vitamin D receptors: primary structures and their
RT      origins";
RL      Proc. Natl. Acad. Sci. U.S.A. 91:11596-11600(1994).
CC      -1- FUNCTION: VDR MEDIATES THE ACTION OF VITAMIN D3 BY CONTROLLING
CC      THE EXPRESSION OF HORMONE SENSITIVE GENES.
CC      -1- SUBCELLULAR LOCATION: NUCLEAR.
CC      -1- SUBCELLULAR LOCATION: TWO FORMS (A AND B) ARE PRODUCED BY USE OF
CC      ALTERNATIVE PRODUCTS.
CC      -1- ALTERNATIVE INITIATION CODONS IN THE SAME READING FRAME.
CC      -1- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
CC      A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
CC      -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.
CC      NR1 SUPERFAMILY.
-----
CC      THIS SWISS-PROT ENTRY IS COPYRIGHT. IT IS PRODUCED THROUGH A COLLABORATION
CC      BETWEEN THE SWISS INSTITUTE OF BIOINFORMATICS AND THE EMBL OUTSTATION
CC      AT THE EUROPEAN BIOINFORMATICS INSTITUTE. THERE ARE NO RESTRICTIONS ON
CC      USE BY NON-PROFIT INSTITUTIONS AS LONG AS ITS CONTENT IS IN NO WAY
CC      MODIFIED AND THIS STATEMENT IS NOT REMOVED. USAGE BY AND FOR COMMERCIAL
CC      ENTITIES REQUIRES A LICENSE AGREEMENT (SEE http://www.isb-sib.ch/announce/
CC      OR SEND AN EMAIL TO license@isb-sib.ch).
CC      EMBL: U12641; AAA56725.1.
CC      HSSP: P03372; 1HCO
CC      INTERPRO: IPR000324;
DR

```

DR INTERPRO: IPR000536; -
 DR INTERPRO: IPR001628; -
 DR PFAM: PF00104; hormone_rec. 1.
 DR PFAM: PF00105; zf-C4; 1.
 DR PRINTS: PR00047; STROIDFINGER.
 DR PRINTS: PR00350; VITAMINDR.
 DR PROSITE: PS00031; NUCLEAR_RECEPTOR; 1.
 DR Receptor; Transcription regulation; DNA-binding; Nuclear protein;
 zinc-finger; Alternative initiation.
 KW CHAIN 1 448 VITAMIN D3 RECEPTOR, ISOFORM A.
 FT CHAIN 26 448 VITAMIN D3 RECEPTOR, ISOFORM B.
 FT INIT_MET 26 26 FOR ISOFORM B.
 FT DNA_BIND 44 109 C4-TYPE ZINC FINGERS (TWO).
 FT ZN_FING 44 64 C4-TYPE.
 FT ZN_FING 80 104 C4-TYPE.
 FT DOMAIN 110 212 HINGE.
 FT DOMAIN 213 448 LIGAND-BINDING.
 SQ SEQUENCE 448 AA; 50668 MW; FCFIFC3DEAEAF3E CRC64;

Query Match 32.7%; Score 815.5; DB 1; Length 448;
 Best Local Similarity 40.9%; Pred. No. 5.2e-58;
 Matches 179; Conservative 72; Mismatches 154; Indels 33; Gaps 10;

QY 56 CEDTE---SVPGKPSVNADE-EVGGPOICRYOCGRATGYHNNVTCGCKGFFRRMKN 111
 DB 16 CEGSLQSSDMETPAVGPPEFRNVPFRICGVGDRATGFHNAMTCGCKGFFRRMKN 75
 QY 112 ARLRPFKRGACETRTKRRQOCRLKCLSGKKKMSDEAVERALLTKRKSER 171
 DB 76 AMFTCPF-SGCKKITKDRRRCQACRLKCVDIGMKKEFILTDEYORKRREMLKREEE 134
 QY 172 TGTPLGVOGLTEORRMIRELDAOMKTFPTTSHEFNRLP-----GYL 217
 DB 135 ALKESLAKR-LSEBQAKYNILAEHHTPTTTSDFKFPFRPKRSTATHSSSV 193
 QY 218 SSGCLPESLQAPSRREAKMSQVRKDCSLKVSQLRGEGSVWNTKPPADSGKEIFS 277
 DB 194 SDFSESNDVFGSDAGAFPEPEPEMFNSMLDSESDSPSMNTELP-----LP 246
 QY 278 LLPHADMSTMGKIIIFAVYISFRDLPIEDQISLKGAFELCQRLFTVNAETGT 337
 DB 247 MLPRLADIVSYIOKIVIFAKMIPGFRDLTAEDQIALKLSAIVIMLSNOSTMEDMS 306
 QY 338 WECGR--LSYCLD-TAGFOOLLLEPMKPFHYMLKQLHEEYVLMQALSPSPDPG 394
 DB 307 WTGSGNDKRYKVDYQAGHSDMLLEPLVRFQVGLAKTINLEEHVILMAICISPPDPG 366
 QY 395 VLQHRVVDQLEQFAITLKSTIECNRPQARHFLFKIMAMTELRSINAGHQ--RLLR 452
 DB 367 VQDSIVSIEDRLSDTLQTYIRCHNPPGSRLLYAKMIQTLADRLSNEHSKQVCL 426
 QY 453 IODIHPF-ATPLMOELFG 469
 DB 427 FQPEHSMQTLPLVLEFG 444

RESULT 5
 VDR_CHICK STANDARD; PRT; 451 AA.
 ID VDR_CHICK 04392;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE VITAMIN D3 RECEPTOR (VDR) (1,25-DIHYDROXYVITAMIN D3 RECEPTOR).
 GN VDR OR N111.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-LEHORN; TISSUE-KIDNEY;

RX MEDLINE-97223363; PubMed-9056239;
 RA Lu Z., Hanson K., Deluca H.F.;
 RT "Cloning and origin of the two forms of chicken vitamin D receptor.";
 RL Arch. Biochem. Biophys. 339:99-106(1997).
 RN [2]
 RP SEQUENCE OF 45-114 FROM N.A.
 RA MEDLINE-87149040; PubMed-3029866;
 RX McDonnell D.P., Mangelsdorf D.J., Pike J.W., Haussler M.R.,
 RA O'Malley B.W.;
 RT "Molecular cloning of complementary DNA encoding the avian receptor
 for vitamin D."
 RL Science 235:1214-1217(1987).
 CC -1- FUNCTION: VDR MEDIATES THE ACTION OF VITAMIN D3 BY CONTROLLING
 CC THE EXPRESSION OF HORMONE SENSITIVE GENES.
 CC -1- SUBCELLULAR LOCATION: NUCLEAR.
 CC -1- ALTERNATIVE PRODUCTS: TWO FORMS (A AND B) ARE PRODUCED BY USE OF
 CC ALTERNATIVE INITIATION CODONS.
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN KIDNEY AND INTESTINE.
 CC -1- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
 CC A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
 CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.
 CC NRI SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb.ch/announce/>
 CC or send an email to license@isb.ch).
 CC -----
 DR EMBL: AF01356; AAB62579.1;
 DR INTERPRO: IPR000324; -
 DR INTERPRO: IPR000536; -
 DR INTERPRO: IPR001628; -
 DR PFAM: PF00104; hormone_rec. 1.
 DR PFAM: PF00105; zf-C4; 1.
 DR PRINTS: PR00047; STROIDFINGER.
 DR PRINTS: PR00350; VITAMINDR.
 DR PROSITE: PS00031; NUCLEAR_RECEPTOR; 1.
 KW Receptor; Transcription regulation; DNA-binding; Nuclear protein;
 KW zinc-finger.
 FT CHAIN 1 451 VITAMIN D3 RECEPTOR, FORM A.
 FT CHAIN 15 451 VITAMIN D3 RECEPTOR, FORM B.
 FT INIT_MET 15 15 FOR FORM B.
 FT DNA_BIND 47 112 C4-TYPE ZINC FINGERS (TWO).
 FT ZN_FING 47 67 C4-TYPE.
 FT ZN_FING 83 107 C4-TYPE.
 FT DOMAIN 113 215 HINGE.
 FT DOMAIN 216 451 LIGAND-BINDING.
 SQ SEQUENCE 451 AA; 51299 MW; 2078B6A6C8DBE5FC CRC64;

Query Match 32.7%; Score 815; DB 1; Length 451;
 Best Local Similarity 40.0%; Pred. No. 5.8e-58;
 Matches 181; Conservative 75; Mismatches 155; Indels 42; Gaps 11;

QY 47 SWNH-----ADVFACEDTESVPGKPSVNADEVEGGPOICRYOCGRATGYHNNMT 96
 DB 7 SWDEQOOSNAVLPDADMDTVAASTSLP-DPAGDFRNV--PRICGVCDRATGFHNAMT 63
 QY 97 CEGCKGFFRRMKNRANRLRCPKRGACETRTKRRQOCARLKLCSGKKKEMMSDEA 156
 DB 64 CEGCKGFFRRMKNRANRLRCPKRGACETRTKRRQOCARLKLCSGKKKEMMSDEA 122
 QY 157 VEERRALLTKRKSEKRTGQPLGVGLTEORRMIRELDAOMKTFPTTSHEFNRLP-- 214
 DB 123 VQKRKMILKKEEBALKSLKPK-LSEBQGVIVDTLEAHKHTPDYTSFNNFRPEYR 181
 QY 215 -----GYLSSGCLPESLQAPSRREAKMSQVRKDCSLKVSQLRGEGSVW 262
 DB 182 SKFSSRMATSSSVVSDFSSSDSDVFGSDAFAPPEPEPEMFNSMLDSESDSPSM 241

```

QY 263 NKKPPADSGGKIFSLPHMADMYMFKGISFAKVISYERDLEIEDOISLKGAFEL 322
DB 242 ILELH-----LPLMLPHLADLVSYISQIKYIGFAKMIPEFRDLTAEDIALKSSAIEV 294
QY 323 COLRENTVNAETGWEGR--LSYCLEL-TAGGFQOLLEPMKLFHMKKLOHEEY 379
DB 295 IMLRSQSTMDMSTGSSNDKFKYSDVYQAGHSMOLLEPLVFGVGLKLNHEEH 354
QY 380 VLMQAIISLSPDPGVLQHRVVDQLOEQFATLKSYIECNRPQAPARFLFKIMAMTEL 439
DB 355 VLMACILSPDPGVDLSVESIQDRSLDLOTYIRCRHPPPSRLLYAKMIOKLADL 414
QY 440 KSNIAQHTQ--RLRLQDIHPF-APPLMOELFG 469
DB 415 RSLNEHSKQYRCLSFQEPESHQPLVLEVFG 447

RESULT 6
VDR_XENIA STANDARD: PRT: 422 AA.
AC 013124;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE VITAMIN D3 RECEPTOR (VDR) (1,25-DIHYDROXYVITAMIN D3 RECEPTOR).
GN VDR OR NR11.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae;
OC Xenopodinae; Xenopus.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-KIDNEY;
RX MEDLINE-97307679; PubMed-9165021;
RA Li Y.C., Bergwitz C., Jueppner H., Demay M.B.;
RT Cloning and characterization of the vitamin D receptor from Xenopus
RT laevis.
RL Endocrinology 138:2347-2353(1997).
CC -1- FUNCTION: VDR MEDIATES THE ACTION OF VITAMIN D3 BY CONTROLLING
CC THE EXPRESSION OF HORMONE SENSITIVE GENES.
CC -1- SUBCELLULAR LOCATION: NUCLEAR.
CC -1- TISSUE SPECIFICITY: DETECTED IN ALL THE TISSUE EXAMINED. HIGHEST
CC LEVEL IN SMALL INTESTINE AND SKIN.
CC -1- DEVELOPMENTAL STAGE: FIRST DETECTED AT STAGE 13. INCREASES
CC GRADUALLY AND PEAKS AT STAGE 57-61 THEN DECREASES TO THE LEVEL
CC SEEN IN ADULT.
CC -1- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
CC A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.
CC NRI SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.1sb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
DR EMBL, U91846; AB58585.1;
DR INTERPRO: IPR000324;
DR INTERPRO: IPR000536;
DR INTERPRO: IPR001628;
DR PFAM: PF00105; zf-C4; 1.
DR PRINTS: PR00047; STROIDFINGER.
DR PRINTS: PR00350; VITAMINDR.
DR PROSITE: PS00031; NUCLEAR_RECEPTOR; 1.
KW Receptor; Transcription regulation; DNA-binding; Nuclear protein;
KW Zinc-finger.
FT DNA_BIND 25 90 C4-TYPE ZINC FINGERS (TWO).
FT ZN_FING 25 45 C4-TYPE.
FT ZN_FING 61 85 C4-TYPE.

```

```

FT DOMAIN 91 188 HINGE.
FT DOMAIN 189 422 LIGAND-BINDING.
SQ SEQUENCE 422 AA; 48188 MW; CBA9F2541FEE9D5 CRC64;

Query Match 32.6%; Score 814; DB 1; Length 422;
Best local similarity 41.1%; Pred. No. 6,4e-58;
Matches 172; Conservative 74; Mismatches 123; Indels 50; Gaps 10;

QY 77 POCROGDKATGTHYNNWMTCEGCKGFFRRAMRNALCPFRKGCETTRTRCQCAC 136
DB 22 PRICGCGDAGTGFHFNAMTCGCKGFFRRSMKRMFTCPF-NGDCRTTKDNRRRCQSC 80
QY 137 RLKCLGSGKAKMINSDEAVERRALIRKKSERTGTOPLVGOTLEORMMIRELMA 196
DB 81 RLKRCYDYGKMKFEILDEYQKROMINKRKESEALKSMRK-ISDQOKMIDILERA 139
QY 197 QMTPTTFSHFKNRFLPGVLSGCELPESLQAPSREAAKMSQVRKDCSLKVSQIQLG 256
DB 140 HRTFTDTYEDFNKFR-----PPVRENVDPEFRITR-----SSVHTQG 178
QY 257 ---EDGSVMYKRPADS-----GGRK-IFSLPHMADMYMFKGIS 295
DB 179 SPESDSVFTSSPDSSEHGFSSALFGQREYSWGKSGELSLPHIADLVSYISQIKIIG 238
QY 296 FAKVISYFRLPIEDQISLKGAFELCOLRENTVNAETGWEGR--LSYCLEL-TAG 352
DB 239 FAKMIFRFLIADQDIALKSVIEVIMRSQSTLSDMSTGSEDFKRYVDVTOA 298
QY 353 GFOQLLEPLKRFHYMLKQLQHEEYVLMQAIISLSPDPGVLQHRVVDQLOEQFATL 412
DB 299 GHNNELEPLKRFQVGLKRLDHEEHVLLMAICILSPDPGVLQHRVVDQLOEQFATL 358
QY 413 KSYIECNRPQAPARFLFKIMAMTELRSINAOHTRLRIQDIHPF---IPLMQELF 468
DB 359 QTYILCKHPGSGRLLYAKMIOKLADRLSLNEHSKQYRSISFLPESHMKLPLMLEVF 417

RESULT 7
VDR_RAT STANDARD: PRT: 423 AA.
AC P13053;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE VITAMIN D3 RECEPTOR (VDR) (1,25-DIHYDROXYVITAMIN D3 RECEPTOR).
GN VDR OR NR11.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-89071726; PubMed-2849110;
RA Burmeister J.K., Wiese R.J., Maeda N., Deluca H.F.;
RT Structure and regulation of the rat 1,25-dihydroxyvitamin D3
RT receptor.
RL Proc. Natl. Acad. Sci. U.S.A. 85:9499-9502(1988).
RN [2]
RP SEQUENCE OF 58-423 FROM N.A.
RX MEDLINE-88124963; PubMed-2829212;
RA Burmeister J.K., Maeda N., Deluca H.F.;
RT Isolation and expression of rat 1,25-dihydroxyvitamin D3 receptor
RT cDNA.
RL Proc. Natl. Acad. Sci. U.S.A. 85:1005-1009(1988).
CC -1- FUNCTION: VDR MEDIATES THE ACTION OF VITAMIN D3 BY CONTROLLING
CC THE EXPRESSION OF HORMONE SENSITIVE GENES.
CC -1- SUBCELLULAR LOCATION: NUCLEAR.
CC -1- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
CC A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.
CC NRI SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 DR EMBL: J04147; AAA4089.1; -
 DR PIR: A31761; A31761.
 DR PIR: A31367; A31367.
 DR HSSP: P03372; IHCO.
 DR TRANSFAC: T00882; -
 DR INTERPRO: IPR000324; -
 DR INTERPRO: IPR000536; -
 DR INTERPRO: IPR001628; -
 DR PIR: PF00104; hormone_rec; 1.
 DR PIR: PF00105; zf-C4; 1.
 DR PRINTS: PR00047; STEROIDFINGER.
 DR PRINTS: PR00350; VITAMINDR.
 DR PROSITE: PS00031; NUCLEAR_RECEPTOR; 1.
 KW Receptor; Transcription regulation; DNA-binding; Nuclear protein;
 KW Zinc-finger; Phosphorylation.
 FT DNA_BIND 24 89 C4-TYPE ZINC FINGERS (TWO).
 FT ZN_FING 24 44 C4-TYPE.
 FT ZN_FING 60 84 C4-TYPE.
 FT DOMAIN 90 187 HINGE.
 FT DOMAIN 188 423 LIGAND-BINDING.
 SQ SEQUENCE 423 AA; 47813 MW; 1A0E519A9DCE990 CRC64;

Query Match 32.1%; Score 801; DB 1; Length 423;
 Best Local Similarity 42.0%; Pred. No. 7e-57;
 Matches 177; Conservative 75; Mismatches 135; Indels 28; Gaps 9;

QY 77 POICRGCGDKATGYHFNWTCGCGKGFRRAMKRNARLRCFRRGACETRTKTRCOAC 136
 DB 21 PRICGCGDKATGYHFNWTCGCGKGFRRAMKRNARLRCFRRGACETRTKTRCOAC 136
 QY 137 RLKRCLESCKMKEMIMDSAEVERRALIKRKSERTGLOPLGVGLTEEQRMIRELMDA 196
 DB 80 RLKRCVDMGKMEKILTDEYORKREMIKREELALDLSRPK-LSEEOQHITAILLDA 138
 QY 197 QMKRTDTTFSHKNRRLP-GVL--SSGCELPESLQPSREAAKMSQYKRLCSIKV--- 250
 DB 139 HKHYDPTFADPRFRPRVMDSTGYSR---PLTSGNSSSSSDLTYTSLDME 194
 QY 251 -----SLQKRGEDGSVWNYKRPADSGKKEIFSLPHMADMTYFKGIISPAVISTFRD 305
 DB 195 PSQGSNMDLNGEDSD-----DPSTLIDSLPSMLPHLADIVSYIQKVIQVFAKMIPEFRD 249
 QY 306 LPIDQISLKGAPELQALRPNTVFNAETGTCGRLSYCLEDT---AGGFQOLLLEPM 362
 DB 250 LTSDQIVLLKSSAIEVIMLSNOSFTWMDMSGDSQDYKYDVDSKAGHTLEPL 309
 QY 363 LKPYHMLKLIQHEEYVIMQALISLSPDRPGVQLQHRVVDQLOEPAITLTKSYTECNRPQ 422
 DB 310 IKFQVGLKMLHHEEHVILMAICIVSPDRGVQDAKIVAEIQRSLNTLQTYIRCHNP 369
 QY 423 PAHPLFKTAMLTLEKSIQAHTQ--RLRIODIHPF-ATPLMOELFG 465
 DB 370 PGSHQIYAKMIQKIALDRLSNEHSKQYRSLSPENSMKLTPLVLEVF 419

RESULT 8
 VDR_MOUSE STANDARD; PRT; 422 AA.
 AC P48281;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE VITAMIN D3 RECEPTOR (VDR) (1,25-DIHYDROXYVITAMIN D3 RECEPTOR).
 GN VDR OR NR111.
 OS Mus musculus (Mouse).

CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-95137405; PubMed-7835717;
 RA Kamel Y., Kawada T., Fukuwatari T., Ono T., Kato S., Sugimoto E.;
 RT "Cloning and sequencing of the gene encoding the mouse vitamin D
 RT receptor";
 RL Gene 152:281-282(1995).

CC -1- FUNCTION: VDR MEDIATES THE ACTION OF VITAMIN D3 BY CONTROLLING
 CC THE EXPRESSION OF HORMONE SENSITIVE GENES.
 CC -1- SUBCELLULAR LOCATION: NUCLEAR.
 CC -1- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
 CC A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
 CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.
 CC NRI SUBFAMILY.

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: D31969; BAA06737.1; -
 DR HSSP: P03372; IHCO.
 DR MGD: MGI:103076; VDR.
 DR INTERPRO: IPR000324; -
 DR INTERPRO: IPR000536; -
 DR INTERPRO: IPR001628; -
 DR PIR: PF00104; hormone_rec; 1.
 DR PIR: PF00105; zf-C4; 1.
 DR PRINTS: PR00047; STEROIDFINGER.
 DR PRINTS: PR00350; VITAMINDR.
 DR PROSITE: PS00031; NUCLEAR_RECEPTOR; 1.
 KW Receptor; Transcription regulation; DNA-binding; Nuclear protein;
 KW Zinc-finger; Phosphorylation.
 FT DNA_BIND 24 89 C4-TYPE ZINC FINGERS (TWO).
 FT ZN_FING 24 44 C4-TYPE.
 FT ZN_FING 60 84 C4-TYPE.
 FT DOMAIN 90 186 HINGE.
 FT DOMAIN 187 422 LIGAND-BINDING.
 FT SEQUENCE 422 AA; 47851 MW; 4704CC8172445732 CRC64;

Query Match 31.8%; Score 792.5; DB 1; Length 422;
 Best Local Similarity 41.8%; Pred. No. 3.4e-56;
 Matches 174; Conservative 74; Mismatches 127; Indels 41; Gaps 11;

QY 77 POICRGCGDKATGYHFNWTCGCGKGFRRAMKRNARLRCFRRGACETRTKTRCOAC 136
 DB 21 PRICGCGDKATGYHFNWTCGCGKGFRRAMKRNARLRCFRRGACETRTKTRCOAC 136
 QY 137 RLKRCLESCKMKEMIMDSAEVERRALIKRKSERTGLOPLGVGLTEEQRMIRELMDA 196
 DB 80 RLKRCVDMGKMEKILTDEYORKREMIKREELALDLSRPK-LSEEOQHITAILLDA 138
 QY 197 QMKRTDTTFSHKNRRLP-GVLSSGCELP-----ESLQPSRE--EAAKMS 239
 DB 139 HKHYDPTFADPRFRPRVMDSTGYSRPTLSFGSDSSNSDLTYTSLDMEPVSFS 198
 QY 240 QYKRCDSLKVSLQKRGEDGSVWNYKRPADSGKKEIFSLPHMADMTYFKGIISPAVISTFRD 299
 DB 199 TM--DL-----NEBGS---DDPSVTLIDSLPSMLPHLADIVSYIQKVIQVFAK 242
 QY 300 ISYFRDLPIEDQISLKGAPELQALRPNTVFNAETGTCGRLSYCLEDT---AGGFQO 355
 DB 243 IGFRLDSDQIVLLKSSAIEVIMLSNOSFTWMDMSGDSQDYKYDVDSKAGHTL 302
 QY 357 LLEPLKTHYMLKLIQHEEYVIMQALISLSPDRPGVQLQHRVVDQLOEPAITLTKSYI 416
 DB 303 ELLEPLKTHYMLKLIQHEEYVIMQALISLSPDRPGVQDAKIVAEIQRSLNTLQTYI 362

QY 417 ECNRPAPHRFLFKIMATLRSINAOHQ--RLRIQDHPF-APPLMOELFG 469
DB 363 RCHNPPGSHOLYAKMIQKLADRLSLNEHSHKQYRSLSFQPNKMLTPLYLEVFG 418

RESULT 9

VDR_BOVIN STANDARD; PRT: 424 AA.
AC 028037;
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE VITAMIN D3 RECEPTOR (VDR) (1,25-DIHYDROXYVITAMIN D3 RECEPTOR).
GN VDR OR NR1L1.
OS Bos taurus (Bovine).
OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
RN (1)
RP MEDLINE-97034797; PubMed-8880453;
RA MEDLINE-97034797; Bosworth B.T., Reinhardt T.A.;
RT "Nucleotide sequence of the bovine vitamin D3 receptor."
RL J. Dairy Sci. 79:1313-1315(1996).
CC -1- FUNCTION: VDR MEDIATES THE ACTION OF VITAMIN D3 BY CONTROLLING
CC THE EXPRESSION OF HORMONE SENSITIVE GENES.
CC -1- SUBCELLULAR LOCATION: NUCLEAR.
CC -1- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
CC A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.
CC NR1 SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC
DR EMBL: U50200; AAB01543.1; -
DR HSSP: P03372; IHQ.
DR INTERPRO: IPR000324; -
DR INTERPRO: IPR000536; -
DR INTERPRO: IPR001628; -
DR PFAM: PF00104; hormone_rec. 1.
DR PFAM: PF00105; zf-C4. 1.
DR PRINTS: PR00047; STROIDFINGER.
DR PRINTS: PR00350; VITAMINDR.
DR PROSITE: PS00031; NUCLEAR_RECEPTOR. 1.
KW Receptor; Transcription regulation; DNA-binding; Nuclear protein;
KW Zinc-finger; Phosphorylation.
FT DNA_BIND 21 86
FT ZN_FING 21 41
FT ZN_FING 57 81
FT ZN_FING 87 188
FT DOMAIN 189 424
FT DOMAIN 189 424
FT SEQUENCE 424 AA; 47957 MW; E9E24926CE38CB7D CMC64;

Query Match 31.6%; Score 789; DB 1; Length 424;
Best Local Similarity 42.2%; Pred. No. 6,5e-56;
Matches 174; Conservative 67; Mismatches 143; Indels 28; Gaps 9;

QY 77 PQCRCGCGKATGYHNNVNTCECKGFFRRAKRNALCPFRKACETTRKTRQCAQC 136
DB 18 PRCGCGKATGYHNNVNTCECKGFFRRAKRNALCPFRKACETTRKTRQCAQC 76
QY 137 RLKRCLESGMKKEMSDAVEREALIRKKKSERTGTQPLGVOGLTEORMMIRELDA 196
DB 77 RLKRCVDSGKMKKEMSDAVEREALIRKKKSERTGTQPLGVOGLTEORMMIRELDA 135

QY 197 QMKTFTTTFHFNKFLRPLGVLSG---CELPESLQAPSEERAKKSVQRKLC----- 246
DB 136 HNKTYDPTSDFCQFRPVRVNDGGGSHPRPNRHPPS--FSGDSSSSCCSDHCITSSDM 193
QY 247 ---SLKVSQLRGEGDSVWYKPPADSGGKEIFSLPHMADMSTYFKGLISAKYISYF 303
DB 194 MDSSFSNLDISEDDSP-----DPSVTELSQSLMPLDLVYSIQKYGAKMIPGR 248
QY 304 RDLPIEDQISLKGAFELCOLRFNFVNAETGTEGCRSLY---CLIEDYAGFOQLLE 360
DB 249 RDLSEQIYVLTSSALEVIMLRNNESTMDKSWTCGNDYKRYSDYTKASHLELIE 308
QY 361 PMKFHYMLKRLQHEEYVLMQALISLSPDRPGVLOHRRVDOLOQSOFAITLKSYTECNR 420
DB 309 PLIKFQVGLKMLNHEEHVLMALICVSPDRPGVDAALIEAIQRLSVTLQTYICRH 368
QY 421 PQRHRLFLFKIMATLRSINAOHQ--RLRIQ-DHPFATPLMOELFG 469
DB 369 PPSHLLYAKMIQKLADRLSLNEHSHKQYRSLSFQPNKMLTPLYLEVFG 420

RESULT 10

VDR_HUMAN STANDARD; PRT: 427 AA.
AC P11473;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE VITAMIN D3 RECEPTOR (VDR) (1,25-DIHYDROXYVITAMIN D3 RECEPTOR).
GN VDR OR NR1L1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
RN (1)
RP MEDLINE-88217887; PubMed-2835767;
RA Baker A.R., McDonnell D.P., Hughes M., Crisp T.M., Mangelsdorf D.J.,
Hussler M.R., Pike J.W., Shine J., O'Malley B.W.;
RT "Cloning and expression of full-length cDNA encoding human vitamin D
RT receptor."
RL Proc. Natl. Acad. Sci. U.S.A. 85:3294-3298(1988).
RN (2)
RP MEDLINE-92379083; PubMed-1324736;
RA Goto H., Chen K.S., Prah J.M., Deluca H.F.;
RT "A single receptor identical with that from intestine/747D cells
RT mediates the action of 1,25-dihydroxyvitamin D-3 in HL-60 cells."
RL Biochim. Biophys. Acta 1132:103-108(1992).
RN (3)
RP SEQUENCE FROM N.A.
RC TISSUE-LENS EPITHELIUM;
RA Rae J.L., Shepard A.R.;
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
RN (4)
RP SEQUENCE FROM N.A.
RC MEDLINE-9735582; PubMed-9212063;
RA Miyamoto K., Kesterson R.A., Yamamoto H., Takekuni Y., Nishiwaki E.,
Tatsumi S., Inoue Y., Morita K., Takeda E., Pike J.W.;
RT "Structural organization of the human vitamin D receptor chromosome 1
RT gene and its promoter."
RL Mol. Endocrinol. 11:1165-1179(1997).
RN (5)
RP SEQUENCE OF 24-90 FROM N.A.
RC TISSUE-PERIPHERAL BLOOD;
RA MEDLINE-91210272; PubMed-1850412;
RA Yu X.-P., Kochanik H., Husterer F.G., Manolagas S.C.;
RT "Vitamin D receptor expression in human lymphocytes. Signal
RT requirements and characterization by western blots and DNA
RT sequencing."
RL J. Biol. Chem. 266:7588-7595(1991).
RN (6)
RP VARIANTS ASP-33 AND GLN-73.
RX MEDLINE-89072761; PubMed-2849209;

FT VARIANT 50 50 R -> Q (IN RICKETS).
 FT VARIANT 73 73 /FTID-VAR_004661.
 FT /FTID-VAR_004662.

Query Match 31.6%; Score 789; DB 1; Length 427;
 Best Local Similarity 42.2%; Pred. No. 6, 5e-56;
 Matches 174; Conservative 67; Mismatches 143; Indels 28; Gaps 9;

QY 77 POLCARCGDKATGYHFNVMTCGCKGFFRRAMRNRLRCPRKAGCETRTKRRQCCAC 136
 DB 21 PRICGCGDRAITGFHNNATCEGCKGFFRRAMRNRLRCPRKAGCETRTKRRQCCAC 79
 QY 137 RLKRCDEGKMKKMSDEAVEERALLIRKKSERTGTPLGVQGLTEQRMMIRELMDA 196
 DB 80 RLKRCVIGMKKREFFILDEYQKRREMLIRKREELAKSLRNK-LSEQOGLIALLDA 138
 QY 197 QMTEPTTSHKFNRLPGVLSG---CELPSLQAPREERAKSQRKDLG----- 246
 DB 139 HNKTPPTSDFCQFPVPYKNDGGSHPSRPSRHTPS--FGDSSSSCSDHCITSSDM 196
 QY 247 ---SLKVSILQREGDSVNNYKRPADSGKEIFSLPHMADSTYMFKGIISFAYISYE 303
 DB 197 MSSSSSNLDLSEEDSD-----DPSVTLSELSQSLMPLADLVSYSGQVIGFAKMTPEF 251
 QY 304 RDLPIEDQISLKGAFELCOLFNTVFNAETGTWEGRLSY---CLEDTAGGFOQLLE 360
 DB 252 RDLTSDQIVLTKSSAIEVYMLRSNFSFTMDMSWTCGNQDYKRVSDYTKAGHSELE 311
 QY 361 PLKFTYMLKQLQHEEYVLMQALISLSPDRPGVQLQHVVDQLOQPAITLTKYTECNR 420
 DB 312 PLKFTYMLKQLQHEEYVLMQALISLSPDRPGVQLQHVVDQLOQPAITLTKYTECNR 371
 QY 421 POPARHFLFKIMAMTELRSINAOHTO--RLIRIQ-DIHPRATPLMOELFG 469
 DB 372 PPGSHLTLAKMIQKLADRLSLNEHSKQYRCLSFQPECSMTPLVLVEFG 423

RESULT 11

NR13_HUMAN STANDARD; PRT; 348 AA.
 AC Q14994;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE ORPHAN NUCLEAR RECEPTOR NR13 (CONSTITUTIVE ANDROSTANE RECEPTOR) (CAR)
 DE (ORPHAN NUCLEAR RECEPTOR M667).
 GN NR13.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=LIVER;
 RA MEDLINE-94158827; PubMed-8114692;
 RA Baes M., Gulick T., Choi H.S., Martinoli M.G., Simha D., Moore D.D.;
 RT "A new orphan member of the nuclear hormone receptor superfamily that
 RT interacts with a subset of retinoic acid response elements.";
 RL Mol. Cell. Biol. 14:1544-1552(1994).
 CC -1- FUNCTION: BINDS AND TRANSACTIVATES THE RETINOIC ACID RESPONSE
 CC ELEMENTS THAT CONTROL EXPRESSION OF THE RETINOIC ACID RECEPTOR
 CC BETA 2 AND ALCOHOL DEHYDROGENASE 3 GENES.
 CC -1- SUBUNIT: HETERODIMER OF NR13 AND RXR.
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
 CC -1- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN LIVER.
 CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.
 CC NRI SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial

CC entitles requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: Z30425; CAAB3016.1; -.
 DR EMBL: 603881; -.
 DR INTERPRO: IPR000324; -.
 DR INTERPRO: IPR000326; -.
 DR INTERPRO: IPR001628; -.
 DR INTERPRO: IPR001728; -.
 DR PFAM: PF00104; "hormone-rec; 1."
 DR PFAM: PF00105; "zf-c4; 1."
 DR PRINTS: PR00047; STROIDFINGER.
 DR PRINTS: PR00350; VITAMINDR.
 DR PRINTS: PR00546; THYROIDHORM.
 DR PROSITE: PS00031; NUCLEAR RECEPTOR; 1.
 KW Receptor; Transcription regulation; DNA-binding; Nuclear protein;
 KW Zinc finger; Activator.
 FT DNA_BIND 11 76 C4-TYPE ZINC FINGERS (TWO).
 FT ZN_FING 11 31 C4-TYPE.
 FT ZN_FING 47 71 C4-TYPE.
 SQ SEQUENCE 348 AA; 39558 MW; 92EAD07D7DCB9DA CRC64;

Query Match 30.8%; Score 768; DB 1; Length 348;
 Best Local Similarity 42.2%; Pred. No. 2, 4e-54;
 Matches 164; Conservative 62; Mismatches 109; Indels 54; Gaps 6;

QY 80 CRVCGDKATGYHFNVMTCGCKGFFRRAMRNRLRCPRKAGCETRTKRRQCCACRLR 139
 DB 11 CVCVGDQATGTFHFNATCEGCKGFFRRAMRNRLRCPRKAGCETRTKRRQCCACRLR 69
 QY 140 RLKRCDEGKMKKMSDEAVEERALLIRKKSERTGTPLGVQGLTEQRMMIRELMDOMK 199
 DB 70 KCLDAGKRCMILSAELALRRKQAGRAQGTPTVQ-----LSKEBELIRLLGHNTR 123
 QY 200 TFDTFHFNKFNRLPGVLSGCELPESLQAPSRREARAKSQRKDLGSLVSLQREGDG 239
 DB 124 HGTMTFQFQVQFRPRAHLFIHQ-PLPTLAP----- 153
 QY 260 SWNNYKRPADSGKEIFSLPHMADSTYMFKGIISFAYISYFRDPIEDQISLKGAA 319
 DB 154 -----VPLVTHFADINTFVLOVTKTKLPVFRSLPIDQISLKGAA 198
 QY 320 FELCOLFNTVFNAETGTWEGRLSYCLEDTAG-GFOQLLEPLMKRYMLKQLQHEE 378
 DB 199 VEIHYVANTFCLQTONFLCGPLRYTIDGARGVPEVELELHFNHGLKRLQLOEPE 258
 QY 379 YVLMQALISLSPDRPGVQLQHVVDQLOQPAITLTKYTECNRPQAPARFLFKIMAMTE 438
 DB 259 YVLMQALISLSPDRPGVQLQHVVDQLOQPAITLTKYTECNRPQAPARFLFKIMAMTE 318
 QY 439 LRSTNAOHTORLRIQDIHPRATPLMOEL 467
 DB 319 LRSINAEYQIOHIGLSAM-MPLQEI 346

RESULT 12

NR13_RAT STANDARD; PRT; 358 AA.
 AC Q9QUS1;
 DT 01-OCT-2000 (Rel. 40, Created)
 DT 01-OCT-2000 (Rel. 40, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE ORPHAN NUCLEAR RECEPTOR NR13 (CONSTITUTIVE ANDROSTANE RECEPTOR)
 DE (CAR).
 GN NR13 OR CAR.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathia; Muridae; Murinae; Rattus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STAIN-WISTAR RYTO, AND FISCHER; TISSUE=LIVER;
 RA Yoshinari K., Sueyoshi T., Moore R., Negishi M.;

RT "Sexually dimorphic nuclear translocation of receptor CAR and
 induction of CYP2B1 gene by phenobarbital in rat livers."
 RT Submitted (MAR-1999) to the EMBL/Genbank/DBJ databases.
 CC -1- FUNCTION: BINDS AND TRANSACTIVATES THE RETINOIC ACID RESPONSE
 ELEMENTS THAT CONTROL EXPRESSION OF THE RETINOIC ACID RECEPTOR
 BETA 2 AND ALCOHOL DEHYDROGENASE 3 GENES (BY SIMILARITY).
 CC -1- SUBUNIT: HETERODIMER OF NR1I3 AND RXR (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.
 CC NRI SUBFAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation
 at the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@sib-sib.ch).
 CC
 CC EMBL: AF133095; AAF22567.1; -
 CC EMBL: AF133094; AAF22566.1; -
 CC INTERPRO: IPR000324; -
 CC INTERPRO: IPR000536; -
 CC INTERPRO: IPR001628; -
 CC INTERPRO: IPR001723; -
 CC INTERPRO: IPR001728; -
 CC PFAM: PF00104; hormone_rec; 1.
 CC PFAM: PF00105; zf-C4; 1.
 CC PRINTS: PRO0047; STEROIDFINGER.
 CC PRINTS: PRO0350; VITAMINDR.
 CC PRINTS: PRO0398; STRDHOMER.
 CC PRINTS: PRO0546; THYROIDHOMR.
 CC PROSITE: PS00031; NUCLEAR RECEPTOR; 1.
 CC Receptor; Transcription regulation; DNA-binding; Nuclear protein;
 CC Zinc-finger; Activator.
 CC ZINC-FINGER: 21 86 C4-TYPE ZINC FINGERS (TWO).
 CC FT ZN_FING 21 41 C4-TYPE.
 CC FT ZN_FING 57 81 C4-TYPE.
 CC SEQUENCE 358 AA; 40922 MW; 13691F49CAD8F1ED CRC64;

Query Match 29.2%; Score 729; DB 1; Length 358;
 Best Local Similarity 40.3%; Pred. No. 3.3e-51;
 Matches 160; Conservative 61; Mismatches 122; Indels 54; Gaps 6;

QY 72 EYVGGPQICRVGDKATGHNFMVTCGCGKFFRRAMRNALRCPPFRKACETIRKTR 131
 DB 13 ESEYGRNCVCGDRATGTHFHALTCGCGKFFRKTVSKTIPICPF-AGRCVSKAQR 71
 QY 132 OCACRLKRCLESGMKKEMISDEAVEERALKRKRSERTGTOPLAGVGLTEQRMIR 191
 DB 72 HCPACGLQCLNVMKMKMILTSALALARRARARRAQKASLQ---8-LSQCKELIQ 125
 QY 192 ELMDAQMKTFTTFFSHFKKFLPGVLSGCELPESLQAFREBAKWSQVKDCLSLVS 251
 DB 126 TLIGATRHVGVPMFDQFVQFRPPAYLFSS-HRPPQPLAV----- 163
 QY 252 LDRGDSVWVKKPPADSGKREISLPHMADSMVYMKGLISPAKYVSPRDLPIEDQ 311
 DB 164 -----VPLPLTHFADINFPVQOLIKTKDLPRLSRSLMEDQ 200
 QY 312 ISLNGAFELCOLRENTVFNATGTWEGRLSYCLEDTAG-GFOOLLLEPMLEPMYMK 370
 DB 201 ISLNGAANEIHLISINTFCLOTQNFQGPCLCYKMDVNHGQYFELLIHHKILK 260
 QY 371 KLQLEHEEYVLMQALISLSPDRPGVQLQHRVVDQLQDQAFITLKSYTECNRPQARHFL 430
 DB 261 RLQLOPEFVALMAAALFSPDRPGVQLQREIIDLQLEVALILNNHIMEQOSRLQSRFLYA 320
 QY 431 KIMAMTELRSTNACHTORLLRIODIHPRATPIQEL 467
 DB 321 KIMGLLAEIRINSASVEIHRIGLSAM-MPLLEGI 356

RESULT 13
 NR1I3 MOUSE STANDARD; PRT; 358 AA.
 AC 035627; 035628;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE ORPHAN NUCLEAR RECEPTOR NR1I3 (CONSTITUTIVE ANDROSTANE RECEPTOR)
 (CAR).
 GN NR1I3 OR CAR.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-LIVER.
 RX MEDLINE-97442417; PubMed-9295294;
 RA Choi H.-S., Chung W., Tzameil I., Simha D., Lee Y.-K., Seol W.,
 RA Moore D.D.;
 RT "Differential transactivation by two isoforms of the orphan nuclear
 hormone receptor CAR."
 RT J. Biol. Chem. 272:23565-23571(1997).
 CC -1- FUNCTION: BINDS AND TRANSACTIVATES THE RETINOIC ACID RESPONSE
 ELEMENTS THAT CONTROL EXPRESSION OF THE RETINOIC ACID RECEPTOR
 BETA 2 AND ALCOHOL DEHYDROGENASE 3 GENES.
 CC -1- SUBUNIT: HETERODIMER OF NR1I3 AND RXR.
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS: CAR1 (SHOWN HERE) AND CAR2;
 ARE PRODUCED BY ALTERNATIVE SPLICING. CAR2 DOES NOT SEEM TO ACT AS
 A TRANSACTIVATOR.
 CC -1- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN LIVER.
 CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.
 CC NRI SUBFAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation
 at the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@sib-sib.ch).
 CC
 CC EMBL: AF009327; AAC53349.1; -
 CC EMBL: AF009328; AAC53350.1; -
 CC MGD: MGI:1346307; NR1I3.
 CC INTERPRO: IPR000324; -
 CC INTERPRO: IPR000536; -
 CC INTERPRO: IPR001628; -
 CC INTERPRO: IPR001723; -
 CC INTERPRO: IPR001728; -
 CC PFAM: PF00104; hormone_rec; 1.
 CC PFAM: PF00105; zf-C4; 1.
 CC PRINTS: PRO0047; STEROIDFINGER.
 CC PRINTS: PRO0350; VITAMINDR.
 CC PRINTS: PRO0398; STRDHOMER.
 CC PRINTS: PRO0546; THYROIDHOMR.
 CC PROSITE: PS00031; NUCLEAR RECEPTOR; 1.
 CC Receptor; Transcription regulation; DNA-binding; Nuclear protein;
 CC Zinc-finger; Activator; Alternative splicing.
 CC ZINC-FINGER: 21 86 C4-TYPE ZINC FINGERS (TWO).
 CC FT ZN_FING 21 41 C4-TYPE.
 CC FT ZN_FING 57 81 C4-TYPE.
 CC FT VARSPLIC 281 286 DRPGV -> GFCMOS (IN ISOFORM CAR2).
 CC FT VARSPLIC 287 358 MISSING (IN ISOFORM CAR2).
 CC SEQUENCE 358 AA; 40894 MW; 4F07730F78CADBC CRC64;

Query Match 29.1%; Score 725; DB 1; Length 358;
 Best Local Similarity 39.3%; Pred. No. 7e-51;
 Matches 157; Conservative 64; Mismatches 124; Indels 54; Gaps 6;

QY 70 ADEEVGPOICRVGDKATGHNFMVTCGCGKFFRRAMRNALRCPPFRKACETIRKTR 129


```

Db 11 ASEEYGPNCVYCGDRAIGHFHALTEGCGFRRVTSITGICPF-AGRCVSKAQ 69
Qy 130 RROCOACRLKRCESGKEMKEMIMSDAVERALTKRKSRGTQPLGVGLREBQM 189
Db 70 RHHCPACRLQCLNCKMCKMKTLSBALALRAAOARAKASIQ-----LWQOREL 123
Qy 190 IRELMADQAKTFDTTFSHPKFRPLGVLSGCCPLPSIADPSREAAKMSQVRKDLCLK 249
Db 124 VOILGATRHVGLPFDQVQFKPPAYL----- 151
Qy 250 VSIQARGDSVWYKPPADSGCKEIFSLPHMADMTYMKGIISPAKYISYRDLPIE 309
Db 152 -----PMHRRPQPRG--PVLPPLTFHADINTFMVOQITKTKDLPFRSLTME 198
Qy 310 DOISLKGAFELCOLRENTVFNATGECRLSYCLEDTA-GGFOOLLLEPMLKFKHYM 368
Db 199 DOISLKGAAVEILHISLNTFCLOTENFCGCLCYKEDVAVHGFQYELFESTLHFKN 258
Qy 369 LKTLQHEEYVLMQALISLSPDRPGVQLQHRVVDQLOEFAITLKSYTECNRPQAHFPL 428
Db 259 LKGHLQEPYVMAAATVAFSPDRPGVQLQREIDQLQEMALILNHHMQOSRLQSRFL 318
Qy 429 FLKIMAMTELRSINAOHTORLLRIODIHPRATPLMOEL 467
Db 319 YAKLGLADLRSTINMAYTELQRELSAM-TPLLGEI 356

RESULT 14
NRH2_HUMAN STANDARD; PRT; 461 AA.
AC P55055;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE OXYSTEROLS RECEPTOR LXR-BETA (LIVER X RECEPTOR BETA) (NUCLEAR ORPHAN RECEPTOR LXR-BETA) (UBIQUITOUSLY-EXPRESSED NUCLEAR RECEPTOR) (NUCLEAR RECEPTOR NER)
GN NR1H2 OR LXR-B OR NUR OR NER.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo; [1]
RN TISSUE FROM N.A.
RC TISSUE-OSTEOSARCOMA;
RX MEDLINE=93011628; PubMed=7926814;
RA Shinar D.M., Endo N., Rutledge S.J., Vogel R., Rodan G.A., Schmidt A.;
RT "NER, a new member of the gene family encoding the human steroid hormone nuclear receptor."
RT Gene 147:273-276(1994)
RL
CC -1- FUNCTION: ORPHAN RECEPTOR. BINDS PREFERENTIALLY TO DOUBLE-STRANDED OLIGONUCLEOTIDE DIRECT REPEATS HAVING THE CONSENSUS HALF-SITE SEQUENCE 5'-AGGTCA-3' AND 4-NT SPACING (DR-4).
CC -1- SUBUNIT: FORMS A HETERODIMER WITH RXR.
CC -1- SUBCELLULAR LOCATION: NUCLEAR (POTENTIAL).
CC -1- TISSUE SPECIFICITY: UBIQUITOUS.
CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.
CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb.ch/announce/ or send an email to license@isb.slb.ch).
CC EMBL: U07132; AAA61783.1; -
CC DR HSSP: P03372; LHCQ.
CC DR MIM: 600380; -
CC DR INTERPRO: IPR000536; -
CC DR INTERPRO: IPR001628; -
CC PFAM: PF00104; hormone_rec; 1.

```

```

DR PFAM: PF00105; zf-C4; 1.
DR PRINTS: P00047; STROIDFINGER.
DR PROSITE: PS00031; NUCLEAR_RECEPTOR; 1.
KW Receptor; Transcription regulation; DNA-binding; Nuclear protein; Zinc-finger.
FT ZINC_BIND 87 154 C4-TYPE ZINC FINGERS (TWO).
FT ZN_FING 87 107 C4-TYPE.
FT ZN_FING 125 149 C4-TYPE.
FT DOMAIN 231 461 LIGAND-BINDING (POTENTIAL).
FT DOMAIN 173 176 POLY-GLN.
FT DOMAIN 191 POLY-SER.
SQ SEQUENCE 461 AA; 51102 MW; 68CE3D9F9BC5C0BE CRC64;

Query Match
Best Local Similarity 20.1%; Score 500.5; DB 1; Length 461;
Matches 134; Conservative 93; Mismatches 153; Indels 99; Gaps 15;

Qy 33 PRGPEANLE-VPRKESWNAADVFHCEDESVGKPSVNADEVGSPQICRGQDKATGYH 91
Db 39 PGSPDDVGVCTDEASACSTDPVTFDPEEPERKKKGPAFMLGHELCRCVCGDASGFH 98
Qy 92 FNVMTCEGCKGFFRRAMRNARLRCFPR-KGACETRTKTRQCAQLRKLESKREM 150
Db 99 YVNLSCGCKGFFRRSVYVGARFYACRGGGTCQMDAFMRRCQCCRLKCKEAGMREOC 158
Qy 151 IMSDEAVERALIRKRSERTGTQ---PLVQG----- 181
Db 159 VLSEEDIRKKK--IRKQOQESQSOSQSPVGPQSSSASPGASPGSEAGSGSGE 216
Qy 182 ---LTEQGMATRELMADQAKTFDTTFHFNKFLRPLGVLSGCCLEPESLOAPREAAKW 238
Db 217 GVQLRAQELMQQLVAAQLOCNKRSF-----DQKTPW 252
Qy 238 SVQRDCLSLAVSLQKREDSVWYKPPADSGKEIFSLPHMADMTYMKGIISPAK 298
Db 253 P-----LGAD-----PQSRDARQFRA--HFEELALISVQELVDFAK 287
Qy 299 VISYFRDPIEDQISLKGAFELCOLRENTVFNATGEC-----GLSTCLSD-TAGG 353
Db 288 QVPGTQLQREDOQLMLASTETELMTETARYNET--SCITFLDFTYSKDDFHRAG 344
Qy 354 FQOLLLEPMKFRHYMLKTLQHEEYVLMQALISLSPDRPGVQLQHRVVDQLOEFOATLK 413
Db 345 LQVERINPIFEFSRAMRLGLDAEYALLININFSADRVPVQSGREVALQGYEVAL 404
Qy 414 SYIEENRQPAHPLFLKIMATLELSINAOHTORL-LRIODIHPRATPLMOELFGI 470
Db 405 SYTRIKRPQDQLR--FPRMLKLVSLRTLSVSHSEOVFALRLD--KKLPPLLSIEMDV 459

RESULT 15
NRH3_MOUSE STANDARD; PRT; 445 AA.
AC Q920Y9;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE OXYSTEROLS RECEPTOR LXR-ALPHA (LIVER X RECEPTOR ALPHA) (NUCLEAR ORPHAN RECEPTOR LXR-ALPHA).
GN NR1H3 OR LXR-A.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclurognath; Muridae; Murinae; Mus; [1]
RN SEQUENCE FROM N.A.
RA Chen Y.E., Horluch M., Dzau V.J.;
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: ORPHAN RECEPTOR. INTERACTION OF LXR ALPHA WITH RXR SHIFTS RXR FROM ITS ROLE AS A SILENT DNA-BINDING PARTNER TO AN ACTIVE LIGAND-BINDING SUBUNIT IN MEDIATING RETINOID RESPONSES THROUGH TARGET GENES DEFINED BY LXRES (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: NUCLEAR (POTENTIAL).

```

CC -! SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.
CC NRI SUBFAMILY.

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/or_send_an_email_to_license@isb-sib.ch).
CC -----

CC EMBL: AF085745; AAD16050.1; -.
CC HSSP: P20393; 1A6Y.
CC INTERPRO: IPR000536; -.
CC INTERPRO: IPR001628; -.
CC PRAM: PR00104; hormone_rec; 1.
CC PRAM: PR00105; zf-C4; 1.
CC PRINTS: PR00047; STROIDFINGER.
CC PROSITE: PS00031; NUCLEAR_RECEPTOR; 1.
CC Receptor; Transcription regulation; DNA-binding; Nuclear protein;
KW Zinc-finger.
FT DNA_BIND 96 161 C4-TYPE ZINC FINGERS (TWO).
FT ZN_BIND 96 116 C4-TYPE.
FT ZN_FING 132 156 C4-TYPE.
SQ SEQUENCE 445 AA; 50417 MW; 1A426DF38D935731 CRC64;

Query Match 20.0%; Score 498.5; DB 1; Length 445;
Best Local Similarity 30.1%; Pred. No. 1.3e-32;
Matches 129; Conservative 85; Mismatches 143; Indels 71; Gaps 13;

QY 57 EDTSEVSGKSVNADDEVGSGQICRVCGDKATGTHFNVTCEGCKGFFRRAMRNRLRC 116
DB 73 EPTELRPQKRKKGPAPKMGELSCVGDASGPHYNVLSCEGCKGFFRRSVTKGANYVC 132
QY 117 PFRGACITRTKTRCOACRLRKLESCKMKEMMSDEAVEERRALIKRKSE----- 170
DB 133 -HSGGCHPMDTYMRKCKOECLRKRCQAGMRECVLSEQIRLKK--LKRQEEQAQATS 189
QY 171 --RTGTPLGVGLTEQRMRIMELMDAOKTDTTFSHFKNRLPGVILSSGCELPESL 227
DB 190 VSPRVSPQVLPQLSPQLGMIEKLVAACQCNRRSFS----- 228
QY 228 QAPSRSEAKKSOVRKDCSLKVSILQRGEDGSVMNTKPPADSGCKEISLPHMADMS 287
DB 229 --DLRVTFW-PIAPD-----PQSRERQQRFA--HFTELAI 260
QY 288 YMFGLISFAFVISTFYRDLPIEDQISILKGAFFELCOLRFTVFN--AETGTWEGRLSY 345
DB 261 VSVQEIYVFAKQPLQFQLSREDQIALKTSALIEVMLETSRRYNPGESEITF-LKDFSY 319
QY 346 CLEDTA--GGFQQLLEPLKTHYMLKTLQLEBEVYLMQAISLSPDRPGVLQHRVYDQ 404
DB 320 NREDFAKAGLQVEFINPFEFSRANNELOLDAEFALLIAISISADPNVQDQLOVERL 379
QY 405 QEQFATLKSTIECNRPQARHFLFKIMAMTELRSINAQHTORL--LRIDIHFPATP 462
DB 380 QHTIYEALHAIVSYNHDP--LMFPRMLKIVSLRTISVSHSEGVFALRLQD--KLLPP 435
QY 463 LMOELFGI 470
DB 436 ILSEIMDV 443

Search completed: February 18, 2001, 13:43:39
Job time: 5859 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 18, 2001, 14:33:07 ; Search time 57.29 seconds

(without alignments)
560.604 Million cell updates/sec

Title: US-09-143-828-4

Perfect score: 2494

Sequence: 1 MVTTRTHHFKESGLRAPAIR.....ODHPAPAPLMQELKGTGGS 473

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 195891 seqs, 6790055 residues

Total number of hits satisfying chosen parameters: 195891

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

SUMMARIES

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Query Match	Length	DB ID	Description
1	991.5	39.8	386	2	thyroid hormone re
2	817	32.8	420	2	vitamin D receptor
3	815.5	32.7	423	2	vitamin D receptor
4	801	32.1	423	2	1,25-dihydroxyvita
5	792.5	31.8	422	2	vitamin D receptor
6	790.5	31.7	425	2	vitamin D receptor
7	789	31.6	427	2	vitamin D receptor
8	789	31.6	427	2	vitamin D receptor
9	768	30.8	427	2	steroid hormone re
10	500.5	20.1	461	2	steroid hormone re
11	496.5	19.9	445	2	steroid hormone re
12	491.5	19.7	445	2	steroid hormone re
13	489	19.6	446	2	steroid hormone re
14	488	19.6	446	2	steroid hormone re
15	463	18.6	446	2	steroid hormone re
16	452.5	18.1	484	2	steroid hormone re
17	450	18.0	369	2	thyroid hormone re
18	450	18.0	369	2	thyroid hormone re
19	449	18.0	373	2	thyroid hormone re
20	448	18.0	373	2	thyroid hormone re
21	447.5	17.9	456	2	thyroid hormone re
22	447	17.9	456	2	thyroid hormone re
23	444.5	17.8	461	2	thyroid hormone re
24	443	17.8	461	2	thyroid hormone re
25	442	17.7	475	2	thyroid hormone re
26	440.5	17.4	464	2	thyroid hormone re
27	434.5	17.4	478	2	thyroid hormone re
28	434	17.4	416	2	thyroid hormone re
29	434	17.4	448	2	thyroid hormone re

ALIGNMENTS

30	432	17.3	444	2	retinoic acid rece
31	429	17.2	579	2	orphan nuclear hor
32	428	17.2	458	2	retinoic acid rece
33	428	17.2	459	2	retinoic acid rece
34	428	17.2	560	2	retinoic acid rece
35	427.5	17.1	576	2	nuclear receptor R
36	427	17.1	408	2	retinoic acid rece
37	425.5	17.1	408	2	thyroid hormone re
38	425	17.0	455	2	retinoic acid rece
39	424.5	17.0	536	2	retinoic acid rece
40	424.5	17.0	614	2	thyroid hormone re
41	424	17.0	458	2	retinoic acid rece
42	423	17.0	452	2	retinoic acid rece
43	423	17.0	578	2	nuclear receptor R
44	422	16.9	448	2	retinoic acid rece
45	422	16.9	955	4	probable transcript

RESULT 1
thyroid hormone receptor homolog - African clawed frog
C:Species: Xenopus laevis (African clawed frog)
C>Date: 20-May-1994 #sequence, revision 26-May-1995 #text, change 20-Sep-1999
C:Accession: S41497, S38486
R:Smith, D.P.; Mason, C.S.; Jones, E.A.; Old, R.W.
Nucleic Acids Res. 22, 66-71, 1994
A>Title: A novel nuclear receptor superfamily member in Xenopus that associates with
A:Reference number: S41497; MID:94173664
A:Status: preliminary; nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1386 <SM>
A:Cross-references: EMBL:X75163; NID:g410517; PIDN:CA53006.1; PID:g410518
C:Superfamily: unassigned erba-related proteins; erba transforming protein homolog;
C:Keywords: zinc finger
F:35-303/Domain: erba transforming protein homolog <ERBA>

Query Match	39.8%	Score 991.5	DB 2	Length 386
Best Local Similarity	49.0%	Pred. No. 8.4e-70		
Matches 204	Conservative 59	Mismatches 104	Indels 49	Gaps 9
57	EDTSVPGKPSVNADEYVGPOICRVGDKATGYHFNVTCEGKGFRRARKNRRLRC	116		
14	EEEDASNSCGTGEDDDGPKICRACGDRATGHERMATCEGKGFRRARKNRRLSC	73		
117	PFRRACETTRKTRKOCACRLKCLSGMKKEMINSDEAVERRALIKR-KSERGTIO	175		
74	PF-QNSCVINKSNRHCQACRLKCLDIGNKELINSDAVEDRRLIKRHLTKLPPT	132		
176	PLVGQGLTEEDRMKIRLMDQAKTFPTTSYKRNRLPVLVSSGCELPDLAPSEEA	235		
133	PPGA-SLTPGQHLTLQVGAHRTKEDENFTFSNFE-----PRT--	171		
236	AKMSQVNRKDCSLKVLQLEGGDSVWNYKRPADSGKEIFSLPHADSTYMFNGIIS	295		
172	-----RSSDPT---QEPDTS--SEAFMLPHISDLVTTMIKGIIS	207		
296	FAKYSYFRDLPIDQISLKGAFELCOLRFTVFAETGTEGRLSYCLEDT-AGGF	354		
208	FAKMLPFKSLDIEDQIALKGSVAEVSFRFTVFSNDINTWCGPFTYDTDMFLAGF	267		
335	QQLLEPLKHYMLKQLQHEEYVLMQALSLFSPRPVQLHRYVDQOQDFATLKS	414		
268	RQLTLEPLVYRHRMKRLINQSEYVAMALSTFASDRPGVCWEKIQKQEHIALTKD	327		
415	YIEGCR-PPQARFLPLFKIMALTELRSINAGHRTORLRIODIHFPATPLMOELFG	469		
328	FLDSGRPSPQNRLLYRKIMCLETETVNDIHKQLLEIWDIOPDPTPLMRVYFG	383		

RESULT 2

JC7229

vitamin D receptor subtype a - Paralelchthys olivaceus

C:Species: Paralelchthys olivaceus
C:Date: 09-Jun-2000 #sequence_revision 09-Jun-2000 #text_change 18-Aug-2000

C:Accession: JC7229

R:Suzuki, T.; Suzuki, N.; Srivastava, A.S.; Kurokawa, T.

A:Title: Identification of cDNAs encoding two subtypes of vitamin D receptor in flounder

A:Reference number: JC7229

A:Accession: JC7229

A:Molecule type: mRNA

A:Residues: 1-420 <STU>

A:Cross-references: DDBJ:AB037673

A:Comment: This receptor is an important factor in calcium homeostasis and bone formation

C:Genetics:

A:Gene: vdr-a

C:Superfamily: unassigned erba-related proteins; erba transforming protein homology

C:Keywords: bone; calcium transport; DNA binding; hormone receptor; intestine; vitamin D

Query Match 32.8%; Score 817; DB 2; Length 420;

Best Local Similarity 43.1%; Pred. No. 3.8e-56;

Matches 179; Conservative 70; Mismatches 144; Indels 22; Gaps 12;

68 VNAD-EEVGGPOICRGDKNATGYHFNVMTCGCKGFFRRAMKRNALRCFRRGACEIT 126

DB 11 VGPDEFDMNAPRIGVCGDKATGFHFNMTCEGCKGFFRRAMKRNALRCFRRGACEIT 69

QY 127 KTRRQACACRLKRCLESGMKNKEMSDAEVRRALIKRKSERTGTPGVGLTEQ 186

DB 70 KDNRRHQAACRLKRCLESGMKNKEMSDAEVRRALIKRKSERTGTPGVGLTEQ 128

QY 187 RMRLMDAOKMTPTDFTTFHKNRFLP--GVLSGCELPESLQADPREEAANWQVRK 243

DB 129 ARMTSLVEAHKHTYDASYSDFSRFPVREGPYTRSRASRAASHLSIDSSDFNSPE 188

QY 244 DLCSLKV---SLQSGEGSVWNNKPPADSGKKEIFSLPMMAMSTYMERGIISFAKYI 300

DB 189 SV-DTKMNFSLMNTYODGA-----SSPDSSEMTKLSLPHLADLVSTISQKIVGFMMKI 243

QY 301 SYFNDLPTEQDLSLKGAFELCOLRNTVNAETGTEG--RLSYCLSD-TAGGFQOL 357

DB 244 PGFDLPTEQDLSLKGAFELCOLRNTVNAETGTEG--RLSYCLSD-TAGGFQOL 303

QY 358 LLEPMKFNHMLKQLQHEEYVLMQAISLSPRPVGLQHRVYDQLQEQPAILTKSYIE 417

DB 304 LLEPMKFNHMLKQLQHEEYVLMQAISLSPRPVGLQHRVYDQLQEQPAILTKSYIE 363

QY 418 CNRPQAHRELFELKIMAMTELRSINOHQ--RLRLIOTDHPF-ATPLMOELFG 469

DB 364 IN--HPCGRLLYAKMIQKRLADRLSLNEHSHKQYSLSPFQHSKQMLPLVLEVF 416

QY 150451

vitamin D receptor isoform A - Japanese quail

C:Species: Coturnix coturnix japonica (Japanese quail)

C:Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 20-Sep-1999

C:Accession: 150451

R:Elaroussi, M.A.; Prah, J.M.; Deluca, H.F.

Proc. Natl. Acad. Sci. U.S.A. 91, 11596-11600, 1994

A:Title: The avian vitamin D receptors: primary structures and their origins.

A:Reference number: 150451; MUID:95062315

A:Accession: 150451

A:Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: mRNA

A:Residues: 1-448 <EIA>

A:Cross-references: EMBL:U12641; NID:9595500; PIDN:AAA56725.1; PID:9595501

C:Superfamily: unassigned erba-related proteins; erba transforming protein homology

C:Keywords: zinc finger

Query Match 32.1%; Score 801; DB 2; Length 423;

Best Local Similarity 42.0%; Pred. No. 6.8e-55;

Matches 172; Conservative 75; Mismatches 135; Indels 28; Gaps 9;

77 POICVCGDKATGYHFNVMTCGCKGFFRRAMKRNALRCFRRGACEITRTRRQAC 136

F:42-362/Domain: erba transforming protein homology <ERBA>

Query Match 32.7%; Score 815.5; DB 2; Length 448;

Best Local Similarity 40.9%; Pred. No. 5.4e-56;

Matches 179; Conservative 72; Mismatches 154; Indels 33; Gaps 10;

QY 56 CEOTE---SVGKPSVNADE-EEVGGPOICRGDKNATGYHFNVMTCGCKGFFRRAMRN 111

DB 16 CEQEQLQSSDKETPAVGTPEFDNAPRIGVCGDKATGFHFNMTCEGCKGFFRRAMRK 75

QY 112 ARLRPFRRGACEITRTRRQACACRLKRCLESGMKNKEMSDAEVRRALIKRKSER 171

DB 76 AMFTCPF-SGDCKITDNRRHQAACRLKRCLESGMKNKEMSDAEVRRALIKRKEE 134

QY 172 TGTQPLGVGLTEQDLSLKGAFELCOLRNTVNAETGTEG--RLSYCLSD-TAGGFQOL 357

DB 135 ALKESLKP-RLEEDQKVINILLEAHKHTYDASYSDFSRFPVREGPYTRSRASRAASHLSIDSSDFNSPE 188

QY 218 SSGCELPESLQADPREEAANWQVRKDLCSLKVSLQSGEGSVWNNKPPADSGKKEIFSLPMMAMSTYMERGIISFAKYI 300

DB 194 SQDFSESDNDFVGSDFAFPEPMEPQMSNLDSESDSPNNILPH-----LP 246

QY 278 LLEPMKFNHMLKQLQHEEYVLMQAISLSPRPVGLQHRVYDQLQEQPAILTKSYIE 417

DB 247 MLPHLADLVSTISQKIVGFMMKI-PGFDLPTEQDLSLKGAFELCOLRNTVNAETGTEG--RLSYCLSD-TAGGFQOL 303

QY 338 WEGGR--LSECLSD-TAGGFQOLLEPMKFNHMLKQLQHEEYVLMQAISLSPRPVGLQHRVYDQLQEQPAILTKSYIE 417

DB 307 WTGGSNDFKRYKSDYQAGHSNDLEPLKRVGVKQKINLLEHEEVYLMQAISLSPRPVGLQHRVYDQLQEQPAILTKSYIE 417

QY 395 VLOHRVYDQLQEQPAILTKSYIECNRPQAHRELFELKIMAMTELRSINOHQ--RLRLIOTDHPF-ATPLMOELFG 469

DB 367 VQDTSVIESIQRLSPTLOTYTRCHRPFGSKLTLAKMIQKRLADRLSLNEHSHKQYSLSPFQHSKQMLPLVLEVF 416

QY 453 IODIHPF-ATPLMOELFG 469

DB 427 FQPEHSMQMLPLVLEVF 444

QY 150451

1,25-dihydroxyvitamin D-3 receptor - rat

C:Species: Rattus norvegicus (Norway rat)

C:Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 20-Sep-1999

C:Accession: A31761; A31367

R:Burnester, J.K.; Wiese, R.J.; Maeda, N.; Deluca, H.F.

Proc. Natl. Acad. Sci. U.S.A. 85, 9499-9502, 1988

A:Title: Structure and regulation of the rat 1,25-dihydroxyvitamin D-3 receptor.

A:Reference number: A31761; MUID:89071726

A:Accession: A31761

A:Molecule type: mRNA

A:Residues: 1-423 <BUD>

A:Cross-references: GB:J04147; GB:J03630; NID:9203956; PIDN:AAA41089.1; PID:9203957

R:Burnester, J.K.; Maeda, N.; Deluca, H.F.

Proc. Natl. Acad. Sci. U.S.A. 85, 1005-1009, 1988

A:Title: Isolation and expression of rat 1,25-dihydroxyvitamin D-3 receptor CDNA.

A:Reference number: A31367; MUID:88124963

A:Accession: A31367

A:Molecule type: mRNA

A:Residues: 1-423 <BUD>

C:Superfamily: unassigned erba-related proteins; erba transforming protein homology

C:Keywords: DNA binding; transcription regulation; zinc finger

F:22-337/Domain: erba transforming protein homology <ERBA>

F:24-44/Region: zinc finger

F:60-84/Region: zinc finger

Query Match 32.1%; Score 801; DB 2; Length 423;

Best Local Similarity 42.0%; Pred. No. 6.8e-55;

Matches 172; Conservative 75; Mismatches 135; Indels 28; Gaps 9;

77 POICVCGDKATGYHFNVMTCGCKGFFRRAMKRNALRCFRRGACEITRTRRQAC 136

Db 124 HNGTFEOPVQFRPAHFIHQ-PLPLAD----- 153
 Qy 260 SYWNYKPPADSGGKEIFSLPHMADMYMKGIISFAKYSYSPRLPIEDQISLKGAA 319
 Db 154 -----VPLVTHPFDINTFMVLYQIKTKRLOPFRSLPIDQISLKGAA 198
 Qy 320 FELCQRYNYFNAETGWECCGLSYCLIEDTAG-GFOQLLEPMKFRMYLKLQHEEE 378
 Db 199 VEICHIVNTTFCLOTQNFCLGRLRYIEDGARVGFVELELFFHGTLRKLOQDEPE 258
 Qy 379 YLMAQALSFSPDRGVQVQHRVQLOQOPATITKSYIECNRRQPAHREFLIMATLE 438
 Db 259 YVLAAMALFSPDRGVQVQHRVQLOQOPATITKSYIECNRRQPAHREFLIMATLE 318
 Qy 439 LRSINAOHTQRLRIODIHPPAPLMOEL 467
 Db 319 LRSINAEAGVQIHIQISLMV-MELLQEI 346

RESULT 10

JC4014
 Steroid hormone-nuclear receptor NER - human
 C:Species: Homo sapiens (man)
 C>Date: 13-Jun-1995 #sequence_revision 14-Jul-1995 #text_change 24-Sep-1999
 C:Accession: JC4014
 R:Shinar, D.M.; Endo, N.; Rutledge, S.J.; Vogel, R.; Rodan, G.A.; Schmidt, A.
 Gene 147, 273-276, 1994
 A:Title: NER, a new member of the gene family encoding the human steroid hormone nuclear
 A:Reference number: JC4014; MID:95011628
 A:Accession: JC4014
 A:Molecule type: mRNA
 A:Residues: 1-461 <SHI>
 A:Cross-references: GB:007132; NID:9641961; PID:AAA61783.1; PID:9641962
 A:Experimental source: osteosarcoma cells SAOS-2/B10
 C:Genetics:
 A:Gene: GDB:UNR
 A:Cross-references: GDB:389570; OMIM:600380
 A:Map position: 19q13.3-19q13.3
 C:Superfamily: unassigned erba-related proteins; erba transforming protein homology
 C:Keywords: steroid hormone receptor
 F:85-381/Domain: erba transforming protein homology <ERBA>
 F:87-154/Domain: DNA binding #status predicted <BIN>

Query Match 20.1%; Score 500.5; DB 2; Length 461;
 Best Local Similarity 28.0%; Pred. No. 2e-31;

Matches 134; Conservative 93; Mismatches 153; Indels 99; Gaps 15;

Qy 33 PRGPEANLE-VRPKESMNHADFVHCEDTESVPGKPSVNADEVGSGPOICRYCGDKATGYH 91
 Db 39 PGPRPDVGTDBASACTDWTVDPEPEPRKKKKGPAPMGLCHELCRVCGDKASGFH 98
 Qy 92 FNVNTECGKGFRRAMKRNARLRCFPR-KGACETRTTRRQCCARLKLKESGMKEM 150
 Db 99 YNVLSECGKGFRRSVYVGARFYACRGCGTCQADATVARRCCQCCRLKKEKMGMEQC 158
 Qy 151 IMSDEAVERBRALLIRKKKERGTQ-PLTQOG----- 181
 Db 159 VLSSEGIKKKK--IRKQOQOESOSQSPVQSGSSASAPGASPGGSEASOGSGEGE 216
 Qy 182 ---LTEEQMMIRELMDQMKTFDTFSHFKNFLRPLGVLSGCELPESLAPSEAKY 238
 Db 217 GVQLTAAQELMIQQLVAAGLOQCNRRFS-----DQPKVTPW 252
 Qy 239 SOVRKDLCLAKYSLQLRGDSVWYKPPADSGGKEIFSLPHMADMYMKGIISFAK 298
 Db 253 P-----LGAD-----POSRARQORFA--HFTELALISVGEIVDFAK 287
 Qy 299 VLSYFNDLPFIEDQISLKGAAFEICQLRNTYFNNAETGWECC---GRLSYCLD-TAGG 353
 Db 288 QVPGFLQGRDQIALKLASTIEMLTARTARKYHET---ECITFLDFTYSKDFHRAG 344

Qy 354 FOQLLEPMKFRMYLKLQHEEYVLMQALISFSPDRGVQVQHRVQLOQOPATITK 413
 Db 345 LQVFNIPFESFAMKRLGDDDEYALLTAIIFADRNVOEPGRVALLQOPVEALL 404
 Qy 414 SYIECNRRQPAHREFLIMATLELRSINAOHTQRL--LRIDIHPPAPLMOELFGI 470
 Db 405 SYTRIKRPOLOLR--FPRMMLKLVSLRTLSVSHSEGVFALRLQD--KKLPILSEIWDV 459

RESULT 11

AS6043
 Steroid hormone receptor-like protein RUD-1 - rat
 C:Species: Rattus norvegicus (Norway rat)
 C>Date: 01-Dec-1995 #sequence_revision 01-Dec-1995 #text_change 24-Sep-1999
 C:Accession: AS6043
 R:Apfel, R.; Benzrock, D.; Lernhardt, E.; Ortiz, M.A.; Salbert, G.; Pfahl, M.
 Mol. Cell. Biol. 14, 7025-7035, 1994
 A:Title: A novel orphan receptor specific for a subset of thyroid hormone-responsive
 A:Reference number: AS6043; MID:95021230
 A:Accession: AS6043
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-445 <APE>
 A:Cross-references: GB:U11685; NID:9555751; PID:AAA3633.1; PID:9555752
 A:Note: authors translated the codon GAG for residue 73 as Ser
 C:Superfamily: unassigned erba-related proteins; erba transforming protein homology
 C:Keywords: steroid hormone receptor; zinc finger
 F:94-365/Domain: erba transforming protein homology <ERBA>

Query Match 19.9%; Score 496.5; DB 2; Length 445;
 Best Local Similarity 30.1%; Pred. No. 4e-31;

Matches 129; Conservative 84; Mismatches 144; Indels 71; Gaps 13;

Qy 57 EDHESVPGKPSVNADEVGSGPOICRYCGDKATGYHFNVTGCEGKGFRRAMKRNARLRC 116
 Db 73 EPELTPORRRKGPAPKMLGNELCSVCGDKASAFHYNVLSCGKGFRRSVINGARYIC 132
 Qy 117 PFRGACETRTTRRQCCARLKLKESGMKEMSDENAEERALKRKSE----- 170
 Db 133 -HSGGCHPDTMRRKQDCRLKCRHAGMECVLSEQIRLKK--LKROEQAOATS 189
 Qy 171 ---RTGTQPLGVQGLTEQRMIRELMDQMKTFDTFSHFKNFLRPLGVLSGCELPESL 227
 Db 190 VSPRVSSPPQVLPQLSPEDLGMIKELVAAQOCNRRSS----- 228
 Qy 228 QAPSRERAKNVOYKDKLSLQVSLQLRGDSVWYKPPADSGGKEIFSLPHMADMY 287
 Db 229 ---DRLRYTPW-PIAPD-----POSREARQORFA--HFTELAI 260
 Qy 288 YNFKGISFAKYSYFRDLPFIEDQISLKGAAFEICQLRNTYFN--AEFTWCEGLSY 345
 Db 261 VSGVELVDPKALPFLDLSREDQIALKLTSAIEVMLETSRRNPSSSITF-LKDFSY 319
 Qy 346 CLEDTA-GGFOQLLEPMKFRMYLKLQHEEYVLMQALISFSPDRGVQVQHRVQLO 404
 Db 320 NNEDEPAKGLQVEFIWPFESRSNNEQLNDAPFALLIAISFSDRPNVQOQLOVERL 379
 Qy 405 QOQFALTKSYIECNRRQPAHREFLIMATLELRSINAOHTQRL--LRIDIHPPAP 462
 Db 380 QVTVEALHAHYVINHP--DRLMFPRLMKLVSLRTLSVSHSEGVFALRLQD--KKLP 435
 Qy 463 LMOELDEGI 470
 Db 436 ILSEIWDV 443

RESULT 12

138975
 nuclear orphan receptor LXR-alpha - human
 C:Species: Homo sapiens (man)
 C>Date: 16-Feb-1996 #sequence_revision 16-Feb-1996 #text_change 20-Sep-1999
 C:Accession: 138975


```

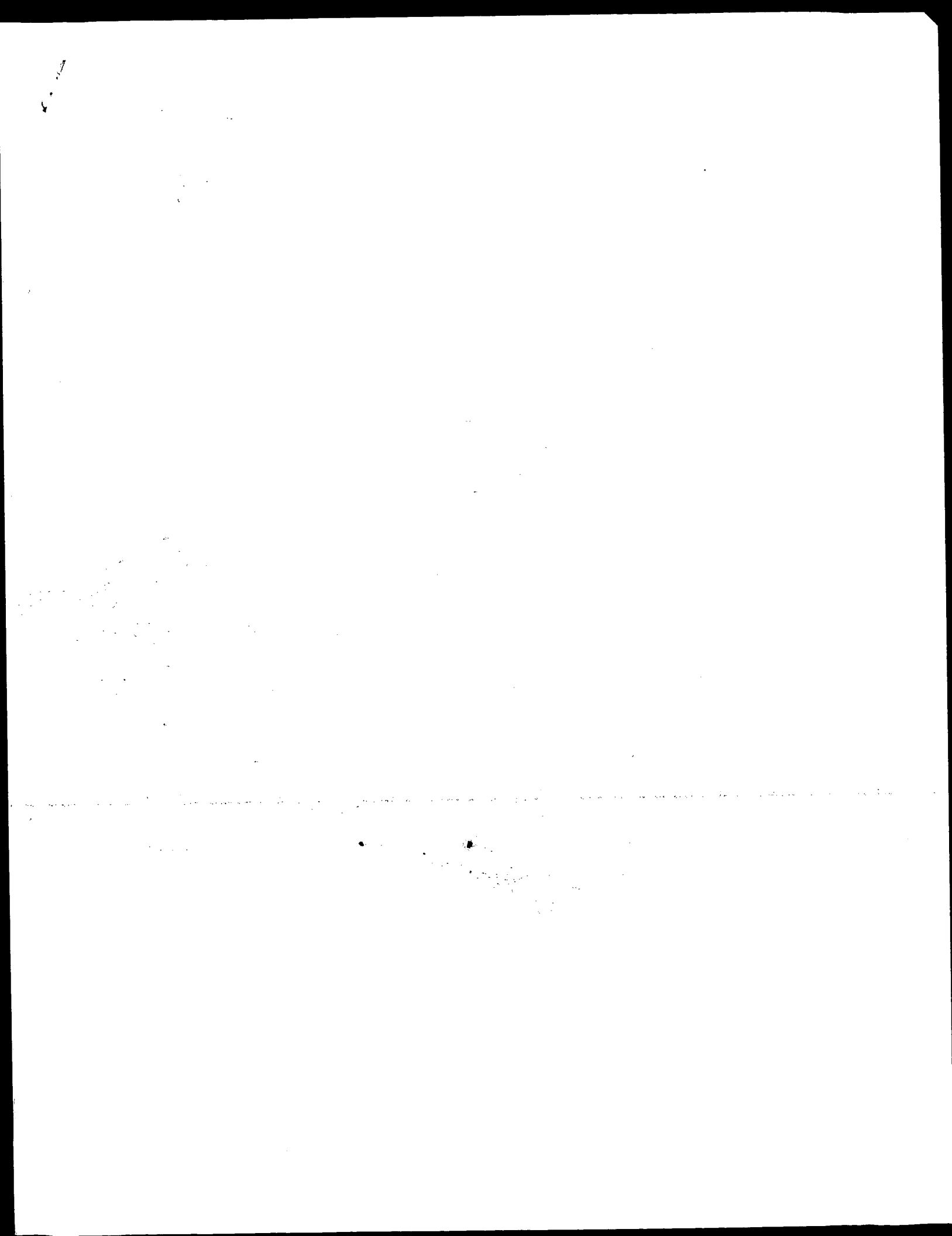
Db      219 EDSEGRDLRCVITSTTKCREKTELVDQOTLLDIDYISKRO-----RKPOEIT 261
Qy      219 SGCELPESLCAPSREANMAQSVOKRDLCSLKVSLQLRGEGDSVMYKPKPADSGGGEISL 276
Db      268 -----NKIKE-----ESAENLI 285
Qy      279 LPHMADSYMFKGIIISFAKVISYVRDLPIEDQISILKGAFLCOLRNTYVFNAGETW 333
Db      284 L---TEATSHQVILVEFKRLPGVQTDHEDQIALKSGAVAMFLRAEELFNKRLPAG 346
Qy      339 ECGSLSTCEEDNAGEFOOLLLEPMIKFHYMLKRLQLEHEEYVLMQASISLFSDFRGVLQ 398
Db      341 HADLLEERIKS--GISDEYIIPMSFEKSYGELKMQOEYALTAIVILSPRDQYIKDR 398
Qy      399 RVVNDQEOEFAILKLSYIECNBPQ-PAHREFYFKIMMLTELRSINAOHTOLL--RQD 455
Db      399 EAVERKQOEPLVDQKLOKCIYQENPQH---FACILGRLELTERTNHHHAELKMSWRKND 455
Qy      456 IHPAPPIWCELGCI 470
Db      456 -HKF-TPPLCEIWDV 468

```

Search completed: February 18, 2001, 14:33:08
Job time: 5357 sec

Oy 456 IHPATPIQELFGI 470
| | | | : : :
Db 456 -HKF-TPILCEINDV 468

Oy 456 IHPATPIQELFGI 470
| | | | : : :
Db 456 -HKF-TPILCEINDV 468



GenCore version 4.5
Copyright (c) 1993 - 2000 Comugen Ltd.

OM protein - protein search, using sw model

Run on: February 18, 2001, 14:31:58 ; Search time 49.49 Seconds
(without alignments)
171.624 Million cell updates/sec

Title: US-09-143-828-4

Perfect score: 2494
Sequence: 1 MYVTRHHEKESLRAPIR.....QDIHPFAPPIWMLGRTGS 473

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 174772 seqs, 17957048 residues

Total number of hits satisfying chosen parameters: 174772

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database:

Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6.COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/PCTUS.COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/Backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	768	30.8	348	1	US-08-459-489-10
2	768	30.8	348	1	US-08-458-686-10
3	768	30.8	348	1	US-07-843-350C-10
4	629	25.2	367	1	US-07-737-736B-4
5	500.5	20.1	461	1	US-08-330-518-2
6	500.5	20.1	461	1	US-08-330-283-2
7	500.5	20.1	461	2	US-08-646-248-2
8	500.5	20.1	461	4	PCT-US95-13924-2
9	500.5	20.1	461	4	PCT-US95-13931-2
10	498	20.0	460	1	US-08-342-411A-2
11	491.5	19.7	440	1	US-08-333-358-8
12	491.5	19.7	440	1	US-08-463-694-8
13	491.5	19.7	440	1	US-08-694-501-8
14	491.5	19.7	447	1	US-08-373-935-1
15	490	19.6	446	2	US-08-372-652-3
16	490	19.6	446	4	PCT-US95-16311-3
17	483.5	19.4	443	1	US-08-342-411A-4
18	472.5	18.9	442	1	US-08-496-631-2
19	461	18.5	469	3	US-08-372-183-2
20	461	18.5	469	4	PCT-US95-17023-2
21	452.5	18.1	484	4	US-08-372-652-1
22	452.5	18.1	484	4	PCT-US95-16311-1
23	445	17.8	451	4	US-08-372-652-2
24	445	17.8	451	4	PCT-US95-16311-2
25	432	17.3	355	5	5223606-4
26	428	17.2	348	5	5223606-2
27	424	17.0	746	3	US-09-144-759-18
28	424	17.0	764	3	US-09-144-759-20

29	418	16.8	410	5	5438126-2	Patent No. 5438126
30	414.5	16.6	433	2	US-08-466-120-2	Sequence 2, Appl1
31	414.5	16.6	433	4	PCT-US94-07266-2	Sequence 2, Appl1
32	413	16.6	462	2	US-08-592-383-2	Sequence 2, Appl1
33	413	16.6	462	2	US-08-095-728B-4	Sequence 4, Appl1
34	413	16.6	462	4	PCT-US92-02320A-4	Sequence 4, Appl1
35	413	16.6	797	4	US-08-095-728B-2	Sequence 2, Appl1
36	413	16.6	797	4	PCT-US92-02320A-2	Sequence 2, Appl1
37	409	16.4	368	5	5223606-3	Patent No. 5223606
38	401.5	16.1	403	2	US-08-592-383-4	Sequence 4, Appl1
39	401	16.1	462	5	5171671-2	Patent No. 5171671
40	392.5	15.7	454	5	5260432-2	Patent No. 5260432
41	371	14.9	475	2	US-08-484-200-2	Sequence 2, Appl1
42	371	14.9	475	3	US-08-465-375-2	Sequence 2, Appl1
43	371	14.9	548	1	US-08-333-358-2	Sequence 2, Appl1
44	371	14.9	548	1	US-08-463-694-2	Sequence 2, Appl1
45	371	14.9	548	1	US-08-694-501-2	Sequence 2, Appl1

ALIGNMENTS

RESULT 1
US-08-459-489-10
Sequence 10, Application US/08459489
Patent No. 5686574
GENERAL INFORMATION:
APPLICANT: David D. Moore et al.
TITLE OF INVENTION: CAR RECEPTORS AND RELATED
TITLE OF INVENTION: MOLECULES AND METHODS
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Fish & Richardson
STREET: 225 Franklin Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 MB
COMPUTER: IBM PS/2 Model 50z or 555X
OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)
SOFTWARE: WordPerfect (Version 5.0)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/459,489
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/843,350
FILING DATE: February 26, 1992
ATTORNEY/AGENT INFORMATION:
NAME: Paul T. Clark
REGISTRATION NUMBER: 30,162
REFERENCE/DOCKET NUMBER: 00786/126001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 348
TYPE: amino acid
STRANDEDNESS: N/A
TOPOLOGY: linear
US-08-459-489-10

Query Match 30.8%; Score 768; DB 1; Length 348;

Best Local Similarity 42.2%; Pred. No. 9.9e-69;

Matches 164; Conservative 62; Mismatches 109; Indels 54; Gaps 6;

QY 80 CRVCGDRAIGHNVMTCECKGFFRRAMKRLRCPFKGACETRTKTRRCCOACRLR 139

Db 11 CVCGDOATGYHFNALTCGCKGFFRRVTSIGPTCPF-AGSCVSKTORRHACRLQ 69
 QY 140 KLESCKMKEMIMSDAVERBALIKRKSEKSTGTPLOVGLTEQRMIMELMDAOK 199
 Db 70 KCLDAGMRKDMILSAEALARRAKOQARRAQTPVO-----LSKQEBELITLLGATRR 123
 QY 200 TFDTSHEKFNRLPGVLSGCELPESLOAPSRREAAKMSQVRKDLCSLKVSLQURGEDG 259
 Db 124 HMGTFEQVQFPPRAHLFIHQ-PLPTLAP----- 153
 QY 260 SWNNTKPPADSGKEIFSLPMDASTYMEKGIISFAKVISYFRDLPIEDQISLKGAA 319
 Db 154 -----VLPLVTHFADINTFMVLQVIKFTKDLPIVFRSLPIEDQISLKGAA 198
 QY 320 FELCOLRNTVYFNAETGWEGRSLCYCLEDTAG-GEQOLLEPMKLFHYMLKQLOHEE 378
 Db 199 VEICHTVNTTFCLOTFONFLCGPLRTTEDGARVGFQVEFELHFFHGTJLRKLOLOPE 258
 QY 379 YVLMQAISLSPDRPGVLOHRVVDLOBOFAITLKSYTECNRPQAPARHFLFKIMAMLTE 438
 Db 259 YVLLAAMALFSPDRPGVLOHRVVDLOBOFAITLKSYTECNRPQAPARHFLFKIMAMLTE 438
 QY 439 LRSINAOHTQRLRIQDIHPATPLMOEL 467
 Db 319 LRSINAEVGYOIHOIGLSAM-MPLQEI 346

RESULT 2

US-08-458-686-10
 ; Sequence 10, Application US/08458686
 ; Patent No. 5710017
 ; GENERAL INFORMATION:
 ; APPLICANT: David D. Moore et al.
 ; TITLE OF INVENTION: CAR RECEPTORS AND RELATED
 ; TITLE OF INVENTION: MOLECULES AND METHODS
 ; NUMBER OF SEQUENCES: 10
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Fish & Richardson
 ; STREET: 225 Franklin Street
 ; CITY: Boston
 ; STATE: Massachusetts
 ; COUNTRY: U.S.A.
 ; ZIP: 02110-2804
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
 ; COMPUTER: IBM PS/2 Model 502 or 55SX
 ; OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)
 ; SOFTWARE: WordPerfect (Version 5.0)
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/458,686
 ; FILING DATE:
 ; CLASSIFICATION: 514
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 07/843,350
 ; FILING DATE: February 26, 1992
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Paul T. Clark
 ; REGISTRATION NUMBER: 30,162
 ; REFERENCE/DOCKET NUMBER: 00786/126001
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (617) 542-5070
 ; TELEFAX: (617) 542-8906
 ; TELEFAX: 200154
 ; INFORMATION FOR SEQ ID NO: 10:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 348
 ; TYPE: amino acid
 ; STRANDEDNESS: N/A
 ; TOPOLOGY: linear
 ; US-08-458-686-10

Query Match

30.8%; Score 768; DB 1; Length 348;

Best Local Similarity 42.2%; Pred No. 9,9e-69;
 Matches 164; Conservative 62; Mismatches 109; Indels 54; Gaps 6;
 QY 80 CVCGDKATGYHFNALTCGCKGFFRRVTSIGPTCPF-AGSCVSKTORRHACRLQ 139
 Db 11 CVCGDOATGYHFNALTCGCKGFFRRVTSIGPTCPF-AGSCVSKTORRHACRLQ 69
 QY 140 KLESCKMKEMIMSDAVERBALIKRKSEKSTGTPLOVGLTEQRMIMELMDAOK 199
 Db 70 KCLDAGMRKDMILSAEALARRAKOQARRAQTPVO-----LSKQEBELITLLGATRR 123
 QY 200 TFDTSHEKFNRLPGVLSGCELPESLOAPSRREAAKMSQVRKDLCSLKVSLQURGEDG 259
 Db 124 HMGTFEQVQFPPRAHLFIHQ-PLPTLAP----- 153
 QY 260 SWNNTKPPADSGKEIFSLPMDASTYMEKGIISFAKVISYFRDLPIEDQISLKGAA 319
 Db 154 -----VLPLVTHFADINTFMVLQVIKFTKDLPIVFRSLPIEDQISLKGAA 198
 QY 320 FELCOLRNTVYFNAETGWEGRSLCYCLEDTAG-GEQOLLEPMKLFHYMLKQLOHEE 378
 Db 199 VEICHTVNTTFCLOTFONFLCGPLRTTEDGARVGFQVEFELHFFHGTJLRKLOLOPE 258
 QY 379 YVLMQAISLSPDRPGVLOHRVVDLOBOFAITLKSYTECNRPQAPARHFLFKIMAMLTE 438
 Db 259 YVLLAAMALFSPDRPGVLOHRVVDLOBOFAITLKSYTECNRPQAPARHFLFKIMAMLTE 438
 QY 439 LRSINAOHTQRLRIQDIHPATPLMOEL 467
 Db 319 LRSINAEVGYOIHOIGLSAM-MPLQEI 346

RESULT 3

US-07-843-350C-10
 ; Sequence 10, Application US/07843350C
 ; Patent No. 5756448
 ; GENERAL INFORMATION:
 ; APPLICANT: David D. Moore et al.
 ; TITLE OF INVENTION: CAR RECEPTORS AND RELATED
 ; TITLE OF INVENTION: MOLECULES AND METHODS
 ; NUMBER OF SEQUENCES: 10
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Fish & Richardson
 ; STREET: 225 Franklin Street
 ; CITY: Boston
 ; STATE: Massachusetts
 ; COUNTRY: U.S.A.
 ; ZIP: 02110-2804
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
 ; COMPUTER: IBM PS/2 Model 502 or 55SX
 ; OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)
 ; SOFTWARE: WordPerfect (Version 5.0)
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/07/843,350C
 ; FILING DATE: February 26, 1992
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER:
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Paul T. Clark
 ; REGISTRATION NUMBER: 30,162
 ; REFERENCE/DOCKET NUMBER: 00786/126001
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (617) 542-5070
 ; TELEFAX: (617) 542-8906
 ; TELEFAX: 200154
 ; INFORMATION FOR SEQ ID NO: 10:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 348
 ; TYPE: amino acid
 ; STRANDEDNESS: N/A

TOPOLOGY: linear
US-07-843-350C-10

Query Match 30.8%; Score 768; DB 1; Length 348;
Best Local Similarity 42.2%; Pred. No. 9,9e-69;
Matches 164; Conservative 62; Mismatches 109; Indels 54; Gaps 6;

80 CVCCKDAGTGHFNVMTCEGCKGFFRRAMKRNARLRCFRRKAGCEITRRKOCACRLR 139
11 CVCCKDAGTGHFNVMTCEGCKGFFRRKAGCEITRRKOCACRLR 69
140 KLEGGKKEKMSDEAVERRALIKKKEKERTGTPGVGGLTEGQRMIRLDAQAK 199
70 KCLAGMKRMDLSNEMALARRAKOORRAQCTPVQ-----LSKEDEIRLGLAHNR 123
200 TFDTFHFHKKRRLPGVLSGCELPESLQAPSRBAKMSQVKKDCLSLVSLQNGEKG 259
124 HMGTFEFQVFRPPLFIHQ-PLPTLP----- 153
260 SVNKKPPADSGKEIFSLPHNADSTYMKGIISFAKVISYRDLPTEDQISLKGA 319
154 -----VLPLTHFADINTFVQLVKTNDLPFRSLPTEQDLSLKGA 198
320 FELQALNFTVNAETGTCGRSLSTCTAG-GRQQLLEPMLKFRYMKKQLHEEE 378
199 VEIHIYLVNTEFLQTNFLCGPLRYTIEDGAVGQVEFLLEHFGHGLRLQLQEP 258
379 VYLMQALSTSPDRPGVQLHRRVVDQLOEPAITKSYIECNRPQARHRLFKIMAMTE 438
259 VYLLAVALSPDRPGVQLHRRVVDQLOEPAITKSYIECNRPQARHRLFKIMAMTE 438
439 LRSINAOHTORLRIODIHPFATPLMOEL 467
319 LRSINAEYGYOIHQIGLSNM-MPLLOEI 346

RESULT 4
US-07-737-736B-4

Sequence 4, Application US/07737736B
Patent No. 5260199

GENERAL INFORMATION:
APPLICANT: Deluca, Hector F.
APPLICANT: Ross, Troy K.
APPLICANT: Prael, Jean M.
TITLE OF INVENTION: Method Of Producing
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Carl R. Schwartz, Esq., c/o Quarles & Brady
STREET: 411 East Wisconsin Avenue
CITY: Milwaukee
STATE: Wisconsin
COUNTRY: U.S.A.
ZIP: 53202

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/737,736B
FILING DATE: 19910730

CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:

NAME: Schwartz, Carl R.

REGISTRATION NUMBER: 29,437

REFERENCE/DOCKET NUMBER: 96-296-2185-2

TELECOMMUNICATION INFORMATION:

TELEPHONE: 414-277-5715

TELEFAX: 414-277-5774

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 367 amino acids
TYPE: AMINO ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Rat.
PUBLICATION INFORMATION:
AUTHORS: Burmester, James K.
AUTHORS: Maeda, No. 5260199yo
AUTHORS: Deluca, Hector F.
TITLE: Isolation and expression of rat
TITLE: 1,25-dihydroxyvitamin D3 receptor cDNA
JOURNAL: Proc. Natl. Acad. Sci. U.S.A.
VOLUME: 85
PAGES: 1005-1009
DATE: February-1988
US-07-737-736B-4

Query Match 25.2%; Score 629; DB 1; Length 367;
Best Local Similarity 38.5%; Pred. No. 8,7e-55;
Matches 144; Conservative 70; Mismatches 132; Indels 28; Gaps 9;

113 RLCPFRKAGCEITRRKTRKOCACRLKCLSGMKRKMIMSDAVERRALIKRKSERT 172
1 RTCPFR-NGCRLTKKRRHRCACRLKRCVDIGMKKEFILTDEYQRKRMIMKKEEA 59
173 GTPLGVGLTEFORMIRELDAQAKTPTDTEFSHFKNRLPGV--SSGCELPESLQAP 230
60 LKDSLRPK-LSERQCHITILDAHKKITDPTADFRRPVRDGGTGSYSPR---P 114
231 SREBAKMSQVRRLCRLV-----SLQNGEDGSVNKKPPADSGKEIFSLPH 281
115 TLISGNSSSSSSDLYTTSLDMWEPGFSNLDNGEDSD----DPSYTLDSLPH 169
282 MADMSTYMKGIISFAKVISYRDLPTEDQISLKGAPELQALNFTVNAETGTCG 341
170 LADLVSTIGKVIYGRKMLPGFRDLTSDQIVTLKSSAIEVIMKRSNOSFTMDMSDCG 229
342 RLSTCLEDT---AGFQOQLLEPMLKFRYMKKQLQHEEVLVQALSTSPDRPGVQLH 398
230 SODKYKDVTVSKAGHTLELEPLIKFQVGLKRLNHEEVLVLAICTVSPDRPGVQDA 289
399 RYVDQLOEPAITKSYIECNRPQARHRLFKIMAMTELRISNAOHTQ--RLRIQDI 456
290 KIVEAIDORTLSTLQTYIRCRHPPGSHQLYAKMIQKLAJLRSINEHSKQYRSLSTQPE 349
457 HPF-ATPLMOELFG 469
350 NSMKLTPVLEVFG 363

RESULT 5
US-08-330-518-2

Sequence 2, Application US/08330518
Patent No. 5607967

GENERAL INFORMATION:

APPLICANT: Friedman, Eitan

APPLICANT: Holloway, M. Katharine

APPLICANT: Rodan, Gideon

APPLICANT: Schmidt, Azriel

APPLICANT: Vogel, Robert

TITLE OF INVENTION: USE OF RECEPTOR POTENTIATORS

NUMBER OF SEQUENCES: 5

CORRESPONDENCE ADDRESS:

ADDRESSEE: Meick & Co., Inc.

STREET: 126 East Lincoln Avenue

CITY: Rahway

STATE: New Jersey

COUNTRY: US

```

      ZIP: 07065-0907
      COMPUTER READABLE FORM:
      MEDIUM TYPE: Floppy disk
      COMPUTER: IBM PC compatible
      OPERATING SYSTEM: PC-DOS/MS-DOS
      SOFTWARE: Patentin Release #1.0, Version #1.25
      CURRENT APPLICATION DATA:
      APPLICATION NUMBER: US/08/330,518
      FILING DATE:
      CLASSIFICATION: 435
      ATTORNEY/AGENT INFORMATION:
      NAME: Dolan, Catherine A.
      REGISTRATION NUMBER: 36,502
      REFERENCE/DOCKET NUMBER: 19316
      TELECOMMUNICATION INFORMATION:
      TELEPHONE: (908) 594-4283
      TELEFAX: (908) 594-4720
      INFORMATION FOR SEQ ID NO: 2:
      SEQUENCE CHARACTERISTICS:
      LENGTH: 461 amino acids
      TYPE: amino acid
      STRANDEDNESS: single
      TOPOLOGY: linear
      MOLECULE TYPE: peptide
      HYPOTHEICAL: NO
      ANTI-SENSE: NO
      US-08-330-518-2

```

```

Query Match      20.1%; Score 500.5; DB 1; Length 461;
Best Local Similarity 28.0%; Pred. No. 8.9e-42;
Matches 134; Conservative 93; Mismatches 153; Indels 99; Gaps 15;

```

```

33 PRGEANLE-VRPKESWNHADFYHCEDETSVPKPSVNADEEVGPQICRVCGRKATGYH 91
| | | : : : | | | : : : | | | : : : | | | : : : | | | : : : |
39 PGGDDPVPVPGDDEASACSTDWVTPDEEPEERKRRKGPAPKMLGHELCRVCGDKASGFH 98
| | | : : : | | | : : : | | | : : : | | | : : : | | | : : : |
92 FNVATCEGCKGFFRRAKRNARLRCPFR-KGACEITRKTROCOARLRCLESQKMKEM 150
| | | : : : | | | : : : | | | : : : | | | : : : | | | : : : |
99 YNVLSCEGCKGFFRRSVYRGARRACRGGGTCOMAFMRKCOQCRKCKEAGNREOC 158
| | | : : : | | | : : : | | | : : : | | | : : : | | | : : : |
151 IMSDEAVERBALIRKKKSERGTQ---PLGVG----- 181
| | | : : : | | | : : : | | | : : : | | | : : : | | | : : : |
159 VLSEQIKKKK--IRKQOQOSQOSQSPVPGSSSSASGPGASGSGSAGEG 216
| | | : : : | | | : : : | | | : : : | | | : : : | | | : : : |
182 ---LTEQRMMIRLEMDAOKMTFTTSHFNKFLPGVLSGCELPESLQAPREBAKW 238
| | | : : : | | | : : : | | | : : : | | | : : : | | | : : : |
217 GVQLAAQELMIQQLVAOQCNKRSPS-----DQPVTPW 252
| | | : : : | | | : : : | | | : : : | | | : : : | | | : : : |
239 SOVRKDLCSLAVSLQLRGEGSVNMYKPPADSGKEIFSLPHMADSTYMGKISFAK 298
| | | : : : | | | : : : | | | : : : | | | : : : | | | : : : |
253 P-----LGAD-----PQSRDARQQRFA--HFTELAIISVQEIYDFAK 287
| | | : : : | | | : : : | | | : : : | | | : : : | | | : : : |
299 VSYFRDLPIDQISLTKGAFFELCOLRFTVNAETGTWEC---GRLSYCLD-TAGG 353
| | | : : : | | | : : : | | | : : : | | | : : : | | | : : : |
288 QVPGFLQGRDQIALKASTIEIMLETARVNHET---ECITPLKDFYSKDDPHRAG 344
| | | : : : | | | : : : | | | : : : | | | : : : | | | : : : |
354 FQOALLPEMLKFNHMKLQJLHEEYVLMQALISFSPDRPGVQLQHRVYDQLOQEFATLK 413
| | | : : : | | | : : : | | | : : : | | | : : : | | | : : : |
345 LQVEFINPFEFSRAMRRIGLDAEYALILAINIFSADRPVQGPGRVREALQOQYVEALL 404
| | | : : : | | | : : : | | | : : : | | | : : : | | | : : : |
414 STTECRPQPARHFLFKIMLMTLRSLNHOHQRL--LRIDIHFAFPIQMELEGI 470
| | | : : : | | | : : : | | | : : : | | | : : : | | | : : : |
405 STTRIKRPDQQR--FPRMLKMLVSLRTLSSVHSEQYVALLQD--KKPLPLSEINDV 459
| | | : : : | | | : : : | | | : : : | | | : : : | | | : : : |

```

```

RESULT 6
US-08-330-283-2
Sequence 2, Application US/08330283
Patent No. 5679518
GENERAL INFORMATION:
APPLICANT: Friedman, Eitan
APPLICANT: Hollaway, M. Katharine

```

```

APPLICANT: Rodan, Gideon
APPLICANT: Rutledge, Su Jane
APPLICANT: Schmidt, Ariel
APPLICANT: Vogel, Robert
TITLE OF INVENTION: METHOD FOR FINDING RECEPTOR POTENTIATORS
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSER: Marck & Co., Inc.
STREET: 126 East Lincoln Avenue
CITY: Rahway
STATE: New Jersey
COUNTRY: US
ZIP: 07065-0907
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/330,283
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Dolan, Catherine A.
REGISTRATION NUMBER: 36,502
REFERENCE/DOCKET NUMBER: 19327
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908) 594-4283
TELEFAX: (908) 594-4720
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 461 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHEICAL: NO
ANTI-SENSE: NO
US-08-330-283-2

```

```

Query Match      20.1%; Score 500.5; DB 1; Length 461;
Best Local Similarity 28.0%; Pred. No. 8.9e-42;
Matches 134; Conservative 93; Mismatches 153; Indels 99; Gaps 15;

```

```

33 PRGEANLE-VRPKESWNHADFYHCEDETSVPKPSVNADEEVGPQICRVCGRKATGYH 91
| | | : : : | | | : : : | | | : : : | | | : : : | | | : : : |
39 PGGDDPVPVPGDDEASACSTDWVTPDEEPEERKRRKGPAPKMLGHELCRVCGDKASGFH 98
| | | : : : | | | : : : | | | : : : | | | : : : | | | : : : |
92 FNVATCEGCKGFFRRAKRNARLRCPFR-KGACEITRKTROCOARLRCLESQKMKEM 150
| | | : : : | | | : : : | | | : : : | | | : : : | | | : : : |
99 YNVLSCEGCKGFFRRSVYRGARRACRGGGTCOMAFMRKCOQCRKCKEAGNREOC 158
| | | : : : | | | : : : | | | : : : | | | : : : | | | : : : |
151 IMSDEAVERBALIRKKKSERGTQ---PLGVG----- 181
| | | : : : | | | : : : | | | : : : | | | : : : | | | : : : |
159 VLSEQIKKKK--IRKQOQOSQOSQSPVPGSSSSASGPGASGSGSAGEG 216
| | | : : : | | | : : : | | | : : : | | | : : : | | | : : : |
182 ---LTEQRMMIRLEMDAOKMTFTTSHFNKFLPGVLSGCELPESLQAPREBAKW 238
| | | : : : | | | : : : | | | : : : | | | : : : | | | : : : |
217 GVQLAAQELMIQQLVAOQCNKRSPS-----DQPVTPW 252
| | | : : : | | | : : : | | | : : : | | | : : : | | | : : : |
239 SOVRKDLCSLAVSLQLRGEGSVNMYKPPADSGKEIFSLPHMADSTYMGKISFAK 298
| | | : : : | | | : : : | | | : : : | | | : : : | | | : : : |
253 P-----LGAD-----PQSRDARQQRFA--HFTELAIISVQEIYDFAK 287
| | | : : : | | | : : : | | | : : : | | | : : : | | | : : : |
299 VSYFRDLPIDQISLTKGAFFELCOLRFTVNAETGTWEC---GRLSYCLD-TAGG 353
| | | : : : | | | : : : | | | : : : | | | : : : | | | : : : |
288 QVPGFLQGRDQIALKASTIEIMLETARVNHET---ECITPLKDFYSKDDPHRAG 344
| | | : : : | | | : : : | | | : : : | | | : : : | | | : : : |
354 FQOALLPEMLKFNHMKLQJLHEEYVLMQALISFSPDRPGVQLQHRVYDQLOQEFATLK 413
| | | : : : | | | : : : | | | : : : | | | : : : | | | : : : |
345 LQVEFINPFEFSRAMRRIGLDAEYALILAINIFSADRPVQGPGRVREALQOQYVEALL 404
| | | : : : | | | : : : | | | : : : | | | : : : | | | : : : |

```

QY 414 SYIECNRPQAPARFLFKIMAMTELRSINAOHTORL--LRIODHPATPLMOELFGI 470
DB 405 SYTRIKRPQDQRL--FPRMLKLVSLRTLSVSHSEOVFALRLQD--KRLPILSEIMDV 459

RESULT 7

US-08-646-248-2

; Sequence 2, Application US/08646248
; Patent No. 5939322
; GENERAL INFORMATION:
; APPLICANT: Friedman, Eitan
; APPLICANT: Holloway, M. Katharine
; APPLICANT: Rodan, Gideon
; APPLICANT: Rutledge, Su Jane
; APPLICANT: Schmidt, Azriel
; APPLICANT: Vogel, Robert
; TITLE OF INVENTION: METHOD FOR FINDING RECEPTOR POTENTIATORS
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merck & Co., Inc.
; STREET: 126 East Lincoln Avenue
; CITY: Rahway
; STATE: New Jersey
; COUNTRY: US
; ZIP: 07065-0907
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/646,248
; FILING DATE: 14-MAY-1996
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/230,283
; FILING DATE: 27-OCT-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Dolan, Catherine A.
; REGISTRATION NUMBER: 36,502
; REFERENCE/DOCKET NUMBER: 19327
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908) 594-4283
; TELEFAX: (908) 594-4720
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 461 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; US-08-646-248-2

Query Match 20.1%; Score 500.5; DB 2; Length 461;
Best Local Similarity 28.0%; Pred. No. 8.9e-42;
Matches 134; Conservative 93; Mismatches 153; Indels 99; Gaps 15;

QY 33 PRGPANLE-VAPKESNMHADFYHCEDESVPKPSVNADEEVGPOICRVCGKATGYH 91
DB 39 PGCPDPVPGTDEASASTVDIPDPPEEPERKRRKGPAPKMLGHELCRVCGDKASGFH 98
QY 92 FVMTCEGKGFFRRAMRNALRCPR-KGACETTRKTRBQCAQCRKKGESGMKREM 150
DB 99 YNVLSCEGKGFFRRSVVGGARRATACGGGTCQADATMRKCCQCRUKCKEAGMRQC 158
QY 151 IMSDAVERRALIRKKSERTGTQ---PLGVQ----- 181
DB 159 VLSEQDIRKK--IRKQOQOSESQSSQSPVPGQSSSSASPGASPGSGEAGSGSGE 216
QY 162 ---LTEBQMMIRELMDQMKTDTTFTHFNKRLPGVLSGCELPESLQAPSRBAKW 238

DB 217 GVQTAQELMTQDLVAADQCKRSF-----DQKVTFW 252
QY 239 SOVRKDCSLKVLQLRGDEGSVMNKKPPADSGCKRIFSLPMDMSTYKGIISFAK 298
DB 253 P-----LGAD-----POSRAKQORFA---HTELATISVOEIVDFAK 287
QY 299 VISYFRDPLIEDQISLKGAFELCOLRENTVNAETGTWEC---GRLSYLED-TAGC 353
DB 288 QVPGFLQGREDDIALKASTIEMLETARYNHET---ECTFLKDTYKDKDFHRAG 344
QY 354 FQOLLEPMKFTYMLKQLQHEEYVLMQALSBSPDPRVQLQHRVQLOQOPATYTK 413
DB 345 LOVEFNPIEFESRRAMRRLGLDAAEYALLAINIFADRPNTVOEPGRVVALQOPYEALL 404
QY 414 SYIECNRPQAPARFLFKIMAMTELRSINAOHTORL--LRIODHPATPLMOELFGI 470
DB 405 SYTRIKRPQDQRL--FPRMLKLVSLRTLSVSHSEOVFALRLQD--KRLPILSEIMDV 459

RESULT 8

PCT-US95-13924-2

; Sequence 2, Application PC/TUS9513924
; GENERAL INFORMATION:
; APPLICANT: Friedman, Eitan
; APPLICANT: Holloway, M. Katharine
; APPLICANT: Rodan, Gideon
; APPLICANT: Rutledge, Su Jane
; APPLICANT: Schmidt, Azriel
; APPLICANT: Vogel, Robert
; TITLE OF INVENTION: METHOD FOR FINDING RECEPTOR POTENTIATORS
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merck & Co., Inc.
; STREET: 126 East Lincoln Avenue
; CITY: Rahway
; STATE: New Jersey
; COUNTRY: US
; ZIP: 07065-0907
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/13924
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Quagliato, Carol S.
; REGISTRATION NUMBER: 35,330
; REFERENCE/DOCKET NUMBER: 19327 PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908) 594-3809
; TELEFAX: (908) 594-4720
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 461 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; PCT-US95-13924-2

Query Match 20.1%; Score 500.5; DB 4; Length 461;
Best Local Similarity 28.0%; Pred. No. 8.9e-42;
Matches 134; Conservative 93; Mismatches 153; Indels 99; Gaps 15;
QY 33 PRGPANLE-VAPKESNMHADFYHCEDESVPKPSVNADEEVGPOICRVCGKATGYH 91

```

Db 39 PGFPDPVPGTDEASASCTDWTVPDPEEPERKKKGPAPKMLGHELCHVCGDKASGFH 98
QY 92 FNVATCEGCKGFRRARAKRNRALRCPPR-KGACEITRTKTRQOACLRCLSSGKMKEM 150
Db 99 YNVLSCEGCKGFRRSVRGARACRGGGTCQMDAPMRKCCQCLRKCKEAGMRQOC 158
QY 151 IMSDEAVERALLTKRKKSERTGTQ---PLGVQO-----181
Db 159 VLSEEDIRKKK--IRKQOQESOSQSPVPGOSSSSASPGASPGSGSAGSGSGE 216
QY 182 ---LTEROQMMIRELMDQMKTFDTFSHFKNFRLPVGLSSGCELPESLOAPSRBEAKY 238
Db 217 GVQVLAQELMIOQLVAQLOCNKRSFS-----DQPKVTPW 252
QY 239 SQVRKDCSLKVSQLRGEGSVNNYKPPADSGKELFSLPHMADSTYFKGIISFAK 298
Db 253 P-----LGAD-----POSBARQORFA--HFTELAIISVQEIYDFAK 287
QY 299 VISYFRDLPIDQISLKGAFELCOLRNFVNAETGTWEC---GLSYCLSD-TAGG 353
Db 288 QVPGFLOGRDQIALKASTIEIMLETARVNHET---ECITFLKDFYTSKDFHRAG 344
QY 354 FOQLLEPMLEKHYMLKKQLQHEEYVLMQALISLSPDRPGVLOHRVVDLOEOFAITLK 413
Db 345 LOVEFINPIFEFSRMRRLGIDDAEYALILAINIFSADRPVQEGREYALQOPEYVALL 404
QY 414 SYIECNRQPAHRFLFKIMAMLELRSINQHTORL--LRIDIHFPATPLMOELFCI 470
Db 405 SYTRIKRPQDQLR--FPRMLKLVSLRTLSSVSHSQVVALRLQD--KLPPLLSIMDV 459

```

RESULT 9

PCT-US95-13931-2

Sequence 2, Application PC/TUS9513931

GENERAL INFORMATION:

APPLICANT: Friedmann, Eitan

APPLICANT: Holloway, M. Katharine

APPLICANT: Rodan, Gideon

APPLICANT: Schmidt, Azriel

APPLICANT: Vogel, Robert

TITLE OF INVENTION: USE OF RECEPTOR POTENTIATORS

NUMBER OF SEQUENCES: 5

CORRESPONDENCE ADDRESS:

ADDRESSEE: Merck & Co., Inc.

STREET: 126 East Lincoln Avenue

CITY: Rahway

STATE: New Jersey

COUNTRY: US

ZIP: 07065-0907

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US95/13931

FILING DATE:

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Quagliato, Carol S.

REGISTRATION NUMBER: 35,330

REFERENCE/DOCKET NUMBER: 19316 PCT

TELECOMMUNICATION INFORMATION:

TELEPHONE: (908) 594-3809

TELEFAX: (908) 594-4720

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 461 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

HYPOTHETICAL: NO

ANTI-SENSE: NO
PCT-US95-13931-2

Query Match 20.1%, Score 500.5; DB 4; Length 461;
Best Local Similarity 28.0%; Pred. No. 8.9e-42;
Matches 134; Conservative 93; Mismatches 153; Indels 99; Gaps 15;

```

QY 33 PRGPEANLE-YPRKSNMHADFEHCEDESVYGRKSVNADSEVGPQICRYCGKATGYH 91
Db 39 PGFPDPVPGTDEASASCTDWTVPDPEEPERKKKGPAPKMLGHELCHVCGDKASGFH 98
QY 92 FNVATCEGCKGFRRARAKRNRALRCPPR-KGACEITRTKTRQOACLRCLSSGKMKEM 150
Db 99 YNVLSCEGCKGFRRSVRGARACRGGGTCQMDAPMRKCCQCLRKCKEAGMRQOC 158
QY 151 IMSDEAVERALLTKRKKSERTGTQ---PLGVQO-----181
Db 159 VLSEEDIRKKK--IRKQOQESOSQSPVPGOSSSSASPGASPGSGSAGSGSGE 216
QY 182 ---LTEROQMMIRELMDQMKTFDTFSHFKNFRLPVGLSSGCELPESLOAPSRBEAKY 238
Db 217 GVQVLAQELMIOQLVAQLOCNKRSFS-----DQPKVTPW 252
QY 239 SQVRKDCSLKVSQLRGEGSVNNYKPPADSGKELFSLPHMADSTYFKGIISFAK 298
Db 253 P-----LGAD-----POSBARQORFA--HFTELAIISVQEIYDFAK 287
QY 299 VISYFRDLPIDQISLKGAFELCOLRNFVNAETGTWEC---GLSYCLSD-TAGG 353
Db 288 QVPGFLOGRDQIALKASTIEIMLETARVNHET---ECITFLKDFYTSKDFHRAG 344
QY 354 FOQLLEPMLEKHYMLKKQLQHEEYVLMQALISLSPDRPGVLOHRVVDLOEOFAITLK 413
Db 345 LOVEFINPIFEFSRMRRLGIDDAEYALILAINIFSADRPVQEGREYALQOPEYVALL 404
QY 414 SYIECNRQPAHRFLFKIMAMLELRSINQHTORL--LRIDIHFPATPLMOELFCI 470
Db 405 SYTRIKRPQDQLR--FPRMLKLVSLRTLSSVSHSQVVALRLQD--KLPPLLSIMDV 459

```

RESULT 10

US-08-342-411A-2

Sequence 2, Application US/08342411A

Patent No. 5639616

GENERAL INFORMATION:

APPLICANT: LIAO, Shutsung

APPLICANT: SONG, Ching

TITLE OF INVENTION: UNIDENTIFIED NUCLEAR RECEPTOR

TITLE OF INVENTION: COMPOSITIONS AND METHODS

NUMBER OF SEQUENCES: 38

CORRESPONDENCE ADDRESS:

ADDRESSEE: Arnold, White & Durkee

STREET: P.O. Box 4433

CITY: Houston

STATE: TX

COUNTRY: USA

ZIP: 77210-4433

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/342,411A

FILING DATE: 18-NOV-1994

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: KITCHELL, BARBARA S.

REGISTRATION NUMBER: 33,928

REFERENCE/DOCKET NUMBER: ARCD154

TELECOMMUNICATION INFORMATION:

TELEPHONE: (512) 418-3000

TELEFAX: (713) 789-2679
 TELE: 79-0924
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 460 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 US-08-342-411A-2

Query Match 20.0%; Score 498; DB 1; Length 460;
 Best Local Similarity 27.4%; Pred. No. 1.6e-41;
 Matches 143; Conservative 87; Mismatches 134; Indels 138; Gaps 18;

20 PLHSAALASLHNPGRPEANLEVRKRESWNHADFHCEDTESVPG--RPSVNADEEVG-- 75
 4 PTSSLDTPLRGNGPQPGRA-----PSSS-----PTVKEBEPWPBGPPDPVGTDEASSA 54
 76 -----GPOICVCGDKATGYHNVMTCEGCKGFFRRAM 108
 55 CSTDWIIPDEEPEPERKKRKPAPMLGHELCVCGDKRSGFHYNVLDCEGCKGFFRRSV 114
 109 KNNARLRCPFR-KGACEITRTROQACRLKCLSGKKKMMINSDEAVEERRALLIRK 167
 115 VGGARRVACRGCGTCQMDAFMRRCQOCRLKCKEAGREGCVLSEQIRKK--IRKQ 172
 168 KSERTGTQ---PLGVG-----LLEQRMIREMD 195
 173 QQQQSGSQSGSVPGQSSSSASGPGASPGSEAGSGEGEVQULNAQDMLAQQLVA 232
 196 AQMTFTDTTFSHFKNRFLPGVLSGCELPESIQADPSREBAKWSQVRKDLCSLKVQLQR 255
 233 AOLQCNKRSFS-----DQPTVTPW-----L 253
 256 GEDGSVWNNKPPADSGKREIFSLPHMADMYMFKGIISFAKVISYRDLPIEDQISL 315
 254 GAD-----PQSRDAKQKRA--HFTELAIISVOEIVDFAKQVGGFQLGREQIALL 303
 316 KGAPELCLRTNTVFNAETGTWEC---GRLSYCLEL-TAGGFQOLLLEPMKPKHLLK 370
 304 KSTIEIMLELTARRNHT---ECITFLKDTYKSDPHRAGLQVEITNPIFEESKMR 360
 371 KIQLEHEEVYLMQAIISLSPDRPGVQKQHRVDOQLQDFATLKSTIECNRPQAPHRFL 430
 361 RGLGDAEYALLAIINIFESADRPVQEPGRVEALQAPYEALSTYRKRRQDQRR--FP 418
 431 KIMAMTELRSINAOHTORL--LRIDIHFPATPLMQELFGI 470
 419 RMLMKLVSLRTLSSVHSEQVAFALRLQD--KKLPILLSIEMDV 458

RESULT 11
 US-08-333-358-8
 Sequence 8, Application US/0833358
 Patent No. 5571696
 GENERAL INFORMATION:
 APPLICANT: EVANS Ph.D., RONALD M.
 APPLICANT: MANGELSDORF Ph.D., DAVID J.
 APPLICANT: ONG Ms., ESTELITA S.
 APPLICANT: ORO Ph.D., ANTHONY E.
 APPLICANT: BORGMEYER Ph.D., UWE K.
 APPLICANT: GIGDERE Ph.D., VINCENT NMN
 APPLICANT: YAO Mr., TSO-PANG NMN
 TITLE OF INVENTION: NOVEL RECEPTORS
 NUMBER OF SEQUENCES: 14
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Pretty, Schroeder, Brueggemann & Clark
 STREET: 444 So. Flower St., Suite 2000
 CITY: Los Angeles
 STATE: CA
 COUNTRY: US
 ZIP: 90071-2921
 COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentia Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/333,358
 FILING DATE:
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/07/761,068
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: Reiter Ph.D., Stephen E.
 REGISTRATION NUMBER: 31192
 REFERENCE/DOCKET NUMBER: P31 8936
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (619) 535-9001
 TELEFAX: (619) 535-8949
 INFORMATION FOR SEQ ID NO: 8:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 440 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-333-358-8

Query Match 19.7%; Score 491.5; DB 1; Length 440;
 Best Local Similarity 29.7%; Pred. No. 6.6e-41;
 Matches 127; Conservative 85; Mismatches 145; Indels 71; Gaps 12;

57 EDIESVPGKPSVNADEEVGPGQICRVCGDKATGYHNVMTCEGCKGFFRRAMRNRLNC 116
 68 EPTFIRPQKRRKGPAPRMGLNELCSVCGDKASGFHYNVLSCEGCKGFFRRSVKGNHYIC 127
 117 PFRKGACEITRTROQACRLKCLSGKKKMMINSDEAVEERRALLIRKSE----- 170
 128 -HSGGHCPMQYMKRCKQCECRLKRCQAGRECVLSEQIRLK--LKRQEEQAHATS 184
 171 ---RTGQPLGVGGLLEQRMIREMDQMKTFDTTFSHFKNRFLPGVLSGCELPESL 227
 185 LPPRRSPQILPOLSPESQGLMTEKLVAQOCCNRRSFS----- 223
 228 QASREBAKWSQVRKDLCSLKVSLQJRGEDGSVWNNKPPADSGKREIFSLPHMADMT 287
 224 ---DLRLVTPWPAW-----DPSRAKQKRA--HFTELAI 255
 288 YMRGIISFAKVISYRDLPIEDQISLKGAEELCOLRENTVFN--AETGTWECGRLSY 345
 256 VSVQEIYDFAKQLPGFQLSREDOIALKTSALIEVMLTSTRYNGSSITF-LKDFSY 314
 346 CLEDTA-GGFQOLLLEPMKPKHLLKQLEHEEVYLMQAIISLSPDRPGVQKQHRVDOQL 404
 315 NREDFAAAGLQVEITNPIFEESKMRNDELNDAEFLLAISIFSADRPVDOQLVERL 374
 405 QDFATLKSTIECNRPQAPHRFLFKIMAMTELRSINAOHTORL--LRIDIHFPATP 462
 375 QHYVEYLHAYVGIHHPH--DLRMFRLMKLVSLRTLSSVHSEQVAFALRLQD--KKLP 430
 463 LMQELFGI 470
 431 LLSIEMDV 438

RESULT 12
 US-08-463-694-8
 Sequence 8, Application US/08463694
 Patent No. 5696233
 GENERAL INFORMATION:
 APPLICANT: EVANS Ph.D., RONALD M.
 APPLICANT: MANGELSDORF Ph.D., DAVID J.
 APPLICANT: ONG Ms., ESTELITA S.
 APPLICANT: ORO Ph.D., ANTHONY E.

Db 224 ---DRLRTVTPWMPAR-----DPHSREARQORRA---HTELA1 255
 QY 288 YMEGIIISFARVSYFEDLPIDQISLKGAAEFCOLRENTVEN--AETGTECGRLSY 345
 Db 256 VSVQEIYDFAKQLPGFQLSREDQIALKTSAEVMLTSRNYNGSSSITF-LKDFSY 314
 QY 346 CLEDTA-GGFOQLLEPMLEKFKHYMLKLOLHEEYVLMQAISLSPDPGVLQHRVVDOL 404
 Db 315 NREDFAKAGLOVERINFIFESFRANMELQINDAEFALLIAISIFSADRPVODLOVERL 374
 QY 405 QEOFAITLKSYTECNRPQAPHRFLFKIMAMLETLSINAQHTORL--LRIDIHFPATP 462
 Db 375 QHTVEALHNAVYSIHHPH--DRLMFPMLMKLVSLRTISSVSEQVAFALRLQD--KKLP 430
 QY 463 LMDELGI 470
 Db 431 LLEIMDV 438

RESULT 14
 US-08-373-935-1
 ; Sequence 1, Application US/08373935
 ; Patent No. 5747661
 ; GENERAL INFORMATION:
 ; APPLICANT: Evans, Ronald M.
 ; APPLICANT: Mangelsdorf, David J.
 ; APPLICANT: Willy, Patricia J.
 ; TITLE OF INVENTION: IDENTIFICATION OF A DISTINCT
 ; NUMBER OF SEQUENCES: 7
 ; RETINOID-RESPONSIVE PATHWAY AND USES THEREFOR
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Pretty Schroeder, Brueggemann & Clark
 ; STREET: 444 South Flower Street, Suite 2000
 ; CITY: Los Angeles
 ; STATE: CA
 ; COUNTRY: USA
 ; ZIP: 90071
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentln Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/373.935
 ; FILING DATE:
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Reiter, Stephen E.
 ; REGISTRATION NUMBER: 31,192
 ; REFERENCE/DOCKET NUMBER: P41 9894
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 619-546-4737
 ; TELEFAX: 619-546-4737
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 447 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-373-935-1

Query Match 19.7%; Score 491.5; DB 1; Length 447;
 Best Local Similarity 29.7%; Pred. No. 6.7e-41;
 Matches 127; Conservative 85; Mismatches 145; Indels 71; Gaps 12;
 QY 57 EDTESVPGKPSVNADEVGGOICVCDKATGYHFNYMTCEGCGFFPRAMKRNARLRC 116
 Db 75 EPTETRPQRRKGPAPKMLGNEICVCGDKASGFHYNTLSCGCKGFFRSVIGKAHYTC 134
 QY 117 PRKGAACETTRTRQCAQACRLKCLSGAKKEMIMSDAEVEERALLIKRKSE----- 170
 Db 135 -HSGCHCPMDIYMRKRCQECRLKRCRQAGMRECEVLSSEOIRLTK--LKRQEEQAHATS 191

QY 171 ---RTGTOPIGVGLTEEQRMIRELMDAQMKTFTDTSHEKFNRLPGVLSGCCLEPSL 227
 Db 192 LPPRRSSPPLILQPSPEQGMTEKLVAAQOCNRRSFS----- 230
 QY 228 QAPSREBAWNSQVRDLCSLKVSLQLRGDEGSVNNYKRPADSGKEIFSLPHADNST 287
 Db 231 ---DRLRTVTPWMPAR-----DPHSREARQORRA---HTELA1 262
 QY 288 YMEGIIISFARVSYFEDLPIDQISLKGAAEFCOLRENTVEN--AETGTECGRLSY 345
 Db 263 VSVQEIYDFAKQLPGFQLSREDQIALKTSAEVMLTSRNYNGSSSITF-LKDFSY 321
 QY 346 CLEDTA-GGFOQLLEPMLEKFKHYMLKLOLHEEYVLMQAISLSPDPGVLQHRVVDOL 404
 Db 322 NREDFAKAGLOVERINFIFESFRANMELQINDAEFALLIAISIFSADRPVODLOVERL 381
 QY 405 QEOFAITLKSYTECNRPQAPHRFLFKIMAMLETLSINAQHTORL--LRIDIHFPATP 462
 Db 382 QHTVEALHNAVYSIHHPH--DRLMFPMLMKLVSLRTISSVSEQVAFALRLQD--KKLP 437
 QY 463 LMDELGI 470
 Db 438 LLEIMDV 445

RESULT 15
 US-08-372-652-3
 ; Sequence 3, Application US/08372652
 ; Patent No. 5932699
 ; GENERAL INFORMATION:
 ; APPLICANT: Moore, David
 ; APPLICANT: Secl, Hongl.
 ; APPLICANT: Chcl, Heng-Sik
 ; TITLE OF INVENTION: RETINOID X RECEPTOR-INTERACTING
 ; NUMBER OF SEQUENCES: 17
 ; POLYPEPTIDES AND RELATED MOLECULES AND METHODS
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Fish & Richardson P.C.
 ; STREET: 225 Franklin Street, Suite 3100
 ; CITY: Boston
 ; STATE: MA
 ; COUNTRY: USA
 ; ZIP: 02110-2804
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentln Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/372.652
 ; FILING DATE: 13-JAN-1995
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Clark, Paul T.
 ; REGISTRATION NUMBER: 30,162
 ; REFERENCE/DOCKET NUMBER: 00786/246001
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 617/542-5070
 ; TELEFAX: 617/542-8906
 ; Telex: 200154
 ; INFORMATION FOR SEQ ID NO: 3:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 446 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: not relevant
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-372-652-3

Query Match 19.6%; Score 490; DB 2; Length 446;
 Best Local Similarity 29.1%; Pred. No. 9.5e-41;
 Matches 125; Conservative 84; Mismatches 128; Indels 92; Gaps 14;

Tue Feb 20 09:10:30 2001

us-09-143-828-4.rai

Page 10

```

0Y 76 GPQICRCVCGDKATGYHHNNWNTCEBCKGFERRAKRNANRLCPFR--KGACEITRKTROCO 134
Db 74 GHEICRCRCVGDASGFPHNVLTSCBCKGFERRSVYHGGAGRIACRGSGTCOMAFMRKCO 133
0Y 135 ACRLRKLCSGEMKEMIMDSDEAVERRALIRKKSER-----TGTOP 176
Db 134 LCRLRKLCSGEMARQCVLSEEQIRKKR--JOKOQOQOPPEPSEPAASSSGRPAASDGE 191
0Y 177 LGVG-----LTEGRNARIRFLMOAKMTPTTSHKRNRLRPGVLSGGELPESLO 228
Db 192 ASSQSGEGEIOITLAAQELMIOQLVAAQLOQNRKSF----- 229
0Y 229 APSREAKMGSQVRKDLCSLKVSLQNGEDGSVNNYPPADSGKELTSLPHMADMTY 288
Db 230 --DQKVPWP-----LGAD-----PQSDANQGRFA--HFTELATI 262
0Y 289 MEKGIISPAKVIYSTRFDLPIDQDLSLKGAFELCOLRNTVFAAEIGTMEC-----GRLS 344
Db 263 SVOCEIVDPAKOVPGFLQGRBDQIALKASTIELMLQTAARIWHEF--ECITFLKDFI 319
0Y 345 YCLD--TAGGFOQLLEPMLKFHWLKKRLQTHHEEYVLMONISLFSDBRQVLOHRYVDO 403
Db 320 YSKDQFHRAGLOVEINIFEFSSAMRLGDGDAEYVLLAIINFSADBRNVOEPSRYEA 379
0Y 404 LOEQAFTLKSYIECNRPQAPHRFLFLKIMAMTELASINAOHQRL--LRIDIHFEAT 461
Db 380 LOQPVEALMLSTRIKRPDOLR--FPMRLMKLVSLTLSSVHSEQVAFRLAD--KKLP 435
0Y 462 PLMOELFGI 470
Db 436 PLUSETMDV 444

```

Search completed: February 18, 2001, 14:31:59
Job time: 6202 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 18, 2001, 14:29:39 ; Search time 63.99 seconds

(without alignments)
252.753 Million cell updates/sec

Title: US-09-143-828-4

Perfect score: 2494

Sequence: 1 MVTYRTHNFKESLRAFAIP.....ODIHPPATPLMKELFGITGS 473

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 268485 seqs, 34193795 residues

Total number of hits satisfying chosen parameters: 268485

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

1: /cgene2_2/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /cgene2_2/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /cgene2_2/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /cgene2_2/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /cgene2_2/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /cgene2_2/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /cgene2_2/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /cgene2_2/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /cgene2_2/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /cgene2_2/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /cgene2_2/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /cgene2_2/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /cgene2_2/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /cgene2_2/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /cgene2_2/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /cgene2_2/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /cgene2_2/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /cgene2_2/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /cgene2_2/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /cgene2_2/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /cgene2_2/gcgdata/geneseq/geneseq/AA2000.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2494	100.0	473	20	Y25411 Human nmr7-1 prote
2	2494	100.0	473	20	Y15936 A human intranucle
3	2494	100.0	473	20	Y15933 A human intranucle
4	2494	100.0	473	20	Y09516 Human vitamin D re
5	2334	93.6	457	20	Y16035 A human intranucle
6	2334	93.6	457	20	Y15932 A human intranucle
7	2334	93.6	457	20	Y25410 Human nmr7 partial
8	2289	91.8	434	20	Y15931 A human intranucle
9	2287	91.7	437	20	Y09515 Human vitamin D re
10	2171	87.0	434	20	Y21799 Human steroid and
11	2167	86.9	414	20	Y42691 Human pregnane X r
12	1585	63.6	316	20	Y42689 H186-pregnane X r

13	979.5	39.3	386	17	R98521
14	801	32.1	423	20	W94623
15	798	32.0	423	19	W47509
16	791	31.7	477	20	Y09035
17	789	31.6	427	20	W68156
18	789	31.6	427	20	Y09054
19	789	31.6	427	20	Y09056
20	768	30.8	430	20	Y09036
21	768	30.8	438	18	W32536
22	763	30.6	348	20	W93902
23	733.5	29.4	357	20	R11346
24	725	29.1	357	20	Y17872
25	656	26.3	358	20	W93903
26	656	26.3	356	19	W37251
27	632	25.3	356	20	W94622
28	500.5	20.1	460	16	R43656
29	500.5	20.1	461	15	R52980
30	500.5	20.1	461	17	R97982
31	500.5	20.1	461	17	R98140
32	498.5	20.0	445	21	Y32374
33	498	20.0	460	18	W25034
34	497.5	19.9	461	17	R66234
35	491.5	19.7	447	17	W03326
36	480	19.6	446	17	R99736
37	488	19.6	446	17	R94169
38	483.5	19.4	443	16	R74739
39	483.5	19.4	443	18	W25035
40	481.5	19.3	440	14	R33744
41	472.5	18.9	472	19	W40072
42	461	18.5	469	17	W03448
43	457.5	18.3	757	20	Y28603
44	455.5	18.3	757	19	W71297
45	452.5	18.1	484	17	R99735

ALIGNMENTS

RESULT	ID	Y25411	standard: Protein; 473 AA.
XX	XX	Y25411;	
XX	XX	06-SEP-1999 (first entry)	
XX	XX	Human nmr7-1 protein.	
XX	XX	Human nmr7-1; nuclear trans-acting receptor protein; human; regulator;	
XX	XX	Identification; downstream target gene; cell proliferation;	
XX	XX	cell development.	
XX	XX	Homo sapiens.	
XX	XX	WO9931129-A1.	
XX	XX	24-JUN-1999.	
XX	XX	11-DEC-1998;	98WO-US26364.
XX	XX	14-OCT-1998;	98US-0104251.
XX	XX	12-DEC-1997;	97US-0069401.
XX	XX	(MERI) MERCK & CO INC.	
XX	XX	Chen F;	
XX	XX	WPI: 1999-405024/34.	
XX	XX	N-PSDB; X78808.	
XX	XX	DNA encoding human nuclear receptors nmr7 and nmr7-1	
XX	XX	Claim 20; Fig 6; 80pp; English.	

Xenopus orphan rec
Rat vitamin D rece
Rat vitamin D rece
Human vitamin D re
Human vitamin D re
Human vitamin D re
Human vitamin D re
Constitutively act
Human CAR receptor
Human CAR receptor
Mouse nuclear rece
Mouse CAR receptor
Rat vitamin D rece
Rat vitamin D rece
Rat vitamin D rece
Human ubiquitinous n
Human recombinant
Human steroid rece
NR receptor poten
Mouse CNRER-1. Mu
Human ubiquitinous n
Human foetal lung
LXR-alpha, orphan
Retinoid X recepto
OR-1 orphan recept
Rat ubiquitinous nuc
Rat ubiquitinous nuc
XR2. Homo sapiens
Human retinoid rec
Farnesoid-activate
EGR polypeptide su
Lucifilla cuprina ec
Retinoid X recepto

XX This invention describes the isolation of the novel human nuclear
 CC receptors nmr7 or nmr7-1. The nmr7 and nmr7-1 proteins are useful in
 CC the identification of downstream target genes and ligands regulating its
 CC activity. The nuclear receptor is involved in the regulation of in vivo
 CC cell proliferation and/or cell development. The nmr7 and nmr7-1
 CC polynucleotides, expression vectors and host cells are useful for the
 CC recombinant production of the protein.

XX Sequence 473 AA;

Query Match 100.0%; Score 2494; DB 20; Length 473;
 Best Local Similarity 100.0%; Pred. No. 7.3e-237;
 Matches 473; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTVTIRTHFFKSGSLRAPAIPLHSAALASNHPRPEANLEVRPKESWNHADVHCEDTE 60
 DB 1 mvtvtrthffksgslrapaiphsaaaelasnhprpeanlevrpkeshnadfvcde 60
 QY 61 SVPGKPSVNADEEVGPOICRVCGDKATGYHFNWTCGCGKGFRRAMKRNARLRCFPRK 120
 DB 61 svpgkpsvnadeevgppqicrvcgdkatgyhfnwtcgckgffrramkknarlrcpfir 120
 QY 121 GACETIRTRKROCCARLRCLESCKKEMIMSDAEVERALIKRKKSERTGTOPGVQ 180
 DB 121 gaceti rtrrqccarlrclcskkmkemsdaeverallikrkksertgtqpvq 180
 QY 181 GLTEEQMMIRLMDAOKTFTDTSFHKFNRLPGVLSGCCLPESLQAPSREBAKWSQ 240
 DB 181 glteeqmmirlemdaoktftdtsfhknrlpgvlsngclpeslqapsrebaakwsq 240
 QY 241 VRKDLCSLKYSLOLRGDEGSVWNYKPPADSGKRTFSLPHMADMSYMKKGITSPAKVI 300
 DB 241 vrkdlsclkyslorgdegsvwnykp padsgkrtfslphmadmsymkkgitstakvi 300
 QY 301 SYFRDLPIEDQISILKGAFAELCOLRENTVFNATGTWECGRLSYCLEDTAGGFQOLLLE 360
 DB 301 syfrdlpi edqisilkgafaelcolrentvfnatgtwecgrlsy cledtaggfqllle 360
 QY 361 PMLKFNHMKKILQHEEYVLMQALISFSPDRPGVLOHRVVDLOQEOFAITLKSIECNR 420
 DB 361 pmkfnhmkkqlqheeyvlmqal isfspdrpgvlohrvvdloqeo faitlksiecnr 420
 QY 421 POPAHRFLFKIMAMTELRSINAOHTORLRIODIHFPATPLMOELFGITGS 473
 DB 421 ppa hrflfkimamtelrsinaohtorlriodihfpatplmoelfgitgs 473

RESULT 2

ID Y15936 standard; Protein; 473 AA.

XX Y15936;

04-AUG-1999 (first entry)

DE A human intranuclear receptor protein.

KW Human; intranuclear receptor protein; drug development; diagnosis;
 treatment.

OS Homo sapiens.

PN JP11127872-A.

PD 18-MAY-1999.

PF 07-AUG-1998; 98JP-0224172.

PR 11-AUG-1997; 97JP-0230335.

PPA (NIBS) JAPAN TOBACCO INC.

XX MPI: 1999-350330/30.
 DR N-PSDB; X59974.

PT New intranuclear receptor protein - useful for drug development and
 diagnosis and treatment of disease

PS Disclosure; Page 32-35; 38pp; Japanese.

CC The present sequence represents a human intranuclear receptor protein.
 CC The nucleic acid sequence was isolated from a human adult cDNA
 CC library using a swedish ANO23 derived probe. The protein can
 CC be used for the development of drugs and diagnosis and treatment
 CC of various diseases.

XX Sequence 473 AA;

Query Match 100.0%; Score 2494; DB 20; Length 473;
 Best Local Similarity 100.0%; Pred. No. 7.3e-237;
 Matches 473; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTVTIRTHFFKSGSLRAPAIPLHSAALASNHPRPEANLEVRPKESWNHADVHCEDTE 60
 DB 1 mvtvtrthffksgslrapaiphsaaaelasnhprpeanlevrpkeshnadfvcde 60
 QY 61 SVPGKPSVNADEEVGPOICRVCGDKATGYHFNWTCGCGKGFRRAMKRNARLRCFPRK 120
 DB 61 svpgkpsvnadeevgppqicrvcgdkatgyhfnwtcgckgffrramkknarlrcpfir 120
 QY 121 GACETIRTRKROCCARLRCLESCKKEMIMSDAEVERALIKRKKSERTGTOPGVQ 180
 DB 121 gaceti rtrrqccarlrclcskkmkemsdaeverallikrkksertgtqpvq 180
 QY 181 GLTEEQMMIRLMDAOKTFTDTSFHKFNRLPGVLSGCCLPESLQAPSREBAKWSQ 240
 DB 181 glteeqmmirlemdaoktftdtsfhknrlpgvlsngclpeslqapsrebaakwsq 240
 QY 241 VRKDLCSLKYSLOLRGDEGSVWNYKPPADSGKRTFSLPHMADMSYMKKGITSPAKVI 300
 DB 241 vrkdlsclkyslorgdegsvwnykp padsgkrtfslphmadmsymkkgitstakvi 300
 QY 301 SYFRDLPIEDQISILKGAFAELCOLRENTVFNATGTWECGRLSYCLEDTAGGFQOLLLE 360
 DB 301 syfrdlpi edqisilkgafaelcolrentvfnatgtwecgrlsy cledtaggfqllle 360
 QY 361 PMLKFNHMKKILQHEEYVLMQALISFSPDRPGVLOHRVVDLOQEOFAITLKSIECNR 420
 DB 361 pmkfnhmkkqlqheeyvlmqal isfspdrpgvlohrvvdloqeo faitlksiecnr 420
 QY 421 POPAHRFLFKIMAMTELRSINAOHTORLRIODIHFPATPLMOELFGITGS 473
 DB 421 ppa hrflfkimamtelrsinaohtorlriodihfpatplmoelfgitgs 473

RESULT 3

ID Y15933 standard; Protein; 473 AA.

XX Y15933;

04-AUG-1999 (first entry)

DE A human intranuclear receptor protein.

KW Human; intranuclear receptor protein; drug development; diagnosis;
 treatment.

OS Homo sapiens.

PN JP11127872-A.

PD 18-MAY-1999.

XX 07-AUG-1998; 98JP-0224172.
 XX 11-AUG-1997; 97JP-0230335.
 XX (NISB) JAPAN TOBACCO INC.
 XX WPI: 1999-350330/30.
 XX N-PSDB: X59968.
 DR New intranuclear receptor protein - useful for drug development and
 DR diagnosis and treatment of disease
 PT Claim 2; Page 17-19; 38pp; Japanese.
 PS The present sequence represents a human intranuclear receptor protein.
 CC The nucleic acid sequence was isolated from a human adult cDNA
 CC library using a wellfish ANO23 derived probe. The protein can
 CC be used for the development of drugs and diagnosis and treatment
 CC of various diseases.
 SO Sequence 473 AA;

Query Match 100.0%; Score 2494; DB 20; Length 473;
 Best Local Similarity 100.0%; Pred. No. 7.3e-237;
 Matches 473; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVTYRTHHREGSLRPAIPALHSAAEELASNHPRGPEANLEVRKSNHADFVCEDE 60
 DB 1 mvttrthhregslrpaipalhsaaeelaalnprgpeanlevrksnhdvfhcedte 60
 QY 61 SVPGKPSVNADEEVGPGQICRVCGDKATGTHFNWTCCEGKGFRRAMRNARLCPFRK 120
 DB 61 svpgkpsvnaadeevgpgqicrvcgdkatgthfnwtccegkgyffrramrnlcpfrk 120
 QY 121 GACEITRTRRQOCACRLKCLSEGGKMKMSDAVEERRALIRKKSERTGTOPLOVQ 180
 DB 121 gacetrtrrrqocacrlkclseggkmmksdaveerralirrkksertgtoplvq 180
 QY 121 gacetrtrrrqocacrlkclseggkmmksdaveerralirrkksertgtoplvq 180
 DB 121 gacetrtrrrqocacrlkclseggkmmksdaveerralirrkksertgtoplvq 180
 QY 181 GTEBORMMIRELMDAOMKFTDTFSHPFNRLPGVLSGCELPESLOAPREBAKXSO 240
 DB 181 gteebormmirelmdaomkftdtfshpfnrlpgvlsgeclpesloaprebaakxso 240
 QY 241 VRKDLCSLKVSQLRGDEGSVWNYKPPADSGKEIFSLPHNADSTYFKGIISFAKYI 300
 DB 241 vrkdlsclkvsqlrgegsvwnykppadsgkeifslphnadstyfkgiisfakyi 300
 QY 301 SYFRDLPIEDQISLKGAEFLCQLRFNTVNAETGTCGRSLCYLEDTAAGFOQLLLE 360
 DB 301 syfrdlpiedqisllkgaeeflcqlrfntvnaetgtwcgrslscyledtagfoqlille 360
 QY 361 PULKFPYMKKIKQLHBEENVLMQVISTSPDRGVLQHRVYDQLOEOBARTLKSTIEENR 420
 DB 361 pulkfpymkkikqlhbeenvlmqvistspdrgvlohryvdydqloebartlkstieenr 420
 QY 421 POPAHRELELKIMAMLETLSINAOHTORLRIQDIHFFAPFLMOELGITGS 473
 DB 421 popahrelelkimamletlsinaohtorlriqdi hffapflmoelgitgs 473
 DB 421 popahrelelkimamletlsinaohtorlriqdi hffapflmoelgitgs 473

RESULT 4
 ID Y09516 standard; Protein; 473 AA.
 AC Y09516;

DT 16-JUL-1999 (first entry)

DE Human vitamin D receptor related gamma 2 protein.

KW Human; vitamin D receptor related protein; VDR; obesity; diabetes;
 KW anorexia; rheumatoid arthritis; lipoprotein defect; hyperlipidaemia;

KW hypercholesterolaemia; hyperlipoproteinaemia; osteoporosis; tumour;
 KW hyperproliferative skin disorder; hyperthyroidism.
 OS Homo sapiens.
 XX W09919354-A1.
 XX 22-APR-1999.
 XX 31-AUG-1998; 98WO-SR01548.
 XX 31-MAR-1998; 98SE-0001148.
 XX 14-OCT-1997; 97SE-0003745.
 XX (PHAA) PHARMACIA & UPJOHN AB.
 PI Berkenstem A, Dahlberg M;
 DR WPI: 1999-302508/25.
 DR N-PSDB: X56243.
 XX New vitamin D receptor related (VDR) polypeptides, useful for
 PT treating obesity, diabetes, anorexia and rheumatoid arthritis
 PS Claim 19; Page 22-24; 35pp; English.

CC The present sequence is a human vitamin D receptor related (VDR)
 CC polypeptide. Human VDR polypeptides and substances which affect VDR
 CC signal transduction, can be used for treating metabolic, proliferative
 CC or inflammatory conditions. They can be used in the manufacture of a
 CC medicament for treating the following conditions: obesity, diabetes,
 CC anorexia, lipoprotein defects, hyperlipidaemia, hypercholesterolaemia or
 CC hyperlipoproteinaemia and osteoporosis, rheumatoid arthritis, benign and
 CC malign tumours, hyperproliferative skin disorders or hyperthyroidism.
 CC Nucleic acid vectors encoding for expression of a VDR polypeptide can
 CC be used for treating metabolic, proliferative or inflammatory conditions,
 CC by introducing them into a mammal. The introduced nucleic acid is then
 CC capable of transforming a cell in vivo and then polypeptide is expressed.
 CC A substance affecting VDR signal transduction, such as an agonist or
 CC antagonist can be used for the manufacture of a medicament for treating
 CC metabolic, proliferative or inflammatory condition.
 CC N.B. The specification specifically claims the VDR nucleic acid and
 CC polypeptide sequences given in figures 1, 4, 7 and 8, but no figures
 CC are given in the specification.

SO Sequence 473 AA;

Query Match 100.0%; Score 2494; DB 20; Length 473;
 Best Local Similarity 100.0%; Pred. No. 7.3e-237;
 Matches 473; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVTYRTHHREGSLRPAIPALHSAAEELASNHPRGPEANLEVRKSNHADFVCEDE 60
 DB 1 mvttrthhregslrpaipalhsaaeelaalnprgpeanlevrksnhdvfhcedte 60
 QY 61 SVPGKPSVNADEEVGPGQICRVCGDKATGTHFNWTCCEGKGFRRAMRNARLCPFRK 120
 DB 61 svpgkpsvnaadeevgpgqicrvcgdkatgthfnwtccegkgyffrramrnlcpfrk 120
 QY 121 GACEITRTRRQOCACRLKCLSEGGKMKMSDAVEERRALIRKKSERTGTOPLOVQ 180
 DB 121 gacetrtrrrqocacrlkclseggkmmksdaveerralirrkksertgtoplvq 180
 QY 121 gacetrtrrrqocacrlkclseggkmmksdaveerralirrkksertgtoplvq 180
 DB 121 gacetrtrrrqocacrlkclseggkmmksdaveerralirrkksertgtoplvq 180
 QY 181 GTEBORMMIRELMDAOMKFTDTFSHPFNRLPGVLSGCELPESLOAPREBAKXSO 240
 DB 181 gteebormmirelmdaomkftdtfshpfnrlpgvlsgeclpesloaprebaakxso 240
 QY 241 VRKDLCSLKVSQLRGDEGSVWNYKPPADSGKEIFSLPHNADSTYFKGIISFAKYI 300
 DB 241 vrkdlsclkvsqlrgegsvwnykppadsgkeifslphnadstyfkgiisfakyi 300
 QY 301 SYFRDLPIEDQISLKGAEFLCQLRFNTVNAETGTCGRSLCYLEDTAAGFOQLLLE 360

```

DB 301 syfdiediedisllkkaafelqrlntvfnctgwegrissyclcdtaggfgqllle 360
QY 361 PMLFHYMKRLQHEEEYVLMQALISLSPDRPGVLOHRVVDQLOEQFAITLKSYIBCNR 420
DB 361 pmkfhymllkllqheeyvlmqalslspdrpgvlgqhrvvdqldgeqfalfklsylecncr 420
QY 421 PQPAHRELFKIMAMLEIRSNQHOTRILRIODIHPPATPLMOELFGITGS 473
DB 421 pqpahrlflkimmamleirsnaghnqrlrlrlqldhplmclpmgelfgitgs 473

RESULT 5
ID Y16035
AC Y16035;
DE 04-AUG-1999 (first entry)
DE A human intranuclear receptor protein.
DE Human; intranuclear receptor protein; drug development; diagnosis;
DE treatment.
DE Homo sapiens.
XX JPI127872-A.
XX 18-MAY-1999.
XX 07-AUG-1998; 98JP-0224172.
XX 11-AUG-1997; 97JP-0230335.
XX (NISB ) JAPAN TOBACCO INC.
XX WPI: 1999-350330/30.
XX N-PSDB: X59975.
XX DR N-PSDB: X59975.
XX PT New intranuclear receptor protein - useful for drug development and
XX diagnosis and treatment of disease
XX PS Disclosure; Page 35-37; 38pp; Japanese.
XX The present sequence represents a human intranuclear receptor protein.
XX CC The nucleic acid sequence was isolated from a human adult cDNA
XX CC library using a swedish ANO23 derived probe. The protein can
XX CC be used for the development of drugs and diagnosis and treatment
XX CC of various diseases.
XX SQ Sequence 457 AA;

Query Match 93.6%; Score 2334; DB 20; Length 457;
Best Local Similarity 98.7%; Pred. No. 3.8e-221;
Matches 443; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
QY 25 AAEIASNHPRGPENLEVRPKESNNHADFYCEPTESYPRGKPSVNADEEVGGPQICRYCG 84
DB 9 aknlppspgrpeanlevrpkesswnhadfvncedtesvpgkpsvnaadeevgpgqicrvcg 68
QY 85 DKATGYHFNVTGCGCKGFFRRAMRNARLRCPPRKGACETTRKTRROCOACRLKCLLES 144
DB 69 dkaatgyhfnvtgckgckgffrramrnarlrcpprkgaacettrktrrrgcacrlrkclles 128
QY 145 GKKKEMINSDEAVERRALIRKKKSERTGTOPLGVOGLTEORMMIRELMDAOKMTPDIT 204
DB 129 gmkkeminsdeaveerrallirkksetrtgtpilvgvgltteqrmrltelmdaqmktidtt 188
QY 205 FSHRKNFPLPVLSGCLPSPSLAPSPREBAKMSQVRKDLCSIKVSLQRLGEGSGSVWNY 264
DB 189 fshkrnflplvlsagclpspslappspreeakmsqvrkdicsikvslqrlgedsgsvwny 248

```

```

QY 265 KPPADSGKEIFSLPLPHADSTYFKGIISFAKVISYFRDLPIEDQISLLKGAFFELCQ 324
DB 249 kppadsgkeifslplphadstymfkglisfakvisyfrdlpiedqisllkgaafelcq 308
QY 325 LRFNTVNAETGTCWCGRLSYCLEBDFAGFQOLLLEPYLKFHYMLKKLQHEEYVLMQA 384
DB 309 lrfntvnaetgtcwcgrlsyclcdtaggfgqlllepmlkfhymllkllqheeyvlmqa 368
QY 385 ISLSPDRPGVLOHRVVDQLOEQFAITLKSYIECNRPQPAHRELFKIMAMLEIRSNQ 444
DB 369 islspdrpgvlgqhrvvdqldgeqfalfklsylecncrpqpahrlflkimmamleirsn 428

QY 445 OHTORLRLRIODIHPPATPLMOELFGITGS 473
DB 429 ohtorlrlriodihppatplmclpmgelfgitgs 457

RESULT 6
ID Y15932
AC Y15932;
DE 04-AUG-1999 (first entry)
DE A human intranuclear receptor protein.
DE Human; intranuclear receptor protein; drug development; diagnosis;
DE treatment.
DE Homo sapiens.
XX JPI127872-A.
XX 18-MAY-1999.
XX 07-AUG-1998; 98JP-0224172.
XX 11-AUG-1997; 97JP-0230335.
XX (NISB ) JAPAN TOBACCO INC.
XX WPI: 1999-350330/30.
XX N-PSDB: X59967.
XX DR N-PSDB: X59967.
XX PT New intranuclear receptor protein - useful for drug development and
XX diagnosis and treatment of disease
XX PS Claim 2; Page 16-17; 38pp; Japanese.
XX The present sequence represents a human intranuclear receptor protein.
XX CC The nucleic acid sequence was isolated from a human adult cDNA
XX CC library using a swedish ANO23 derived probe. The protein can
XX CC be used for the development of drugs and diagnosis and treatment
XX CC of various diseases.
XX SQ Sequence 457 AA;

Query Match 93.6%; Score 2334; DB 20; Length 457;
Best Local Similarity 98.7%; Pred. No. 3.8e-221;
Matches 443; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
QY 25 AAEIASNHPRGPENLEVRPKESNNHADFYCEPTESYPRGKPSVNADEEVGGPQICRYCG 84
DB 9 aknlppspgrpeanlevrpkesswnhadfvncedtesvpgkpsvnaadeevgpgqicrvcg 68
QY 85 DKATGYHFNVTGCGCKGFFRRAMRNARLRCPPRKGACETTRKTRROCOACRLKCLLES 144
DB 69 dkaatgyhfnvtgckgckgffrramrnarlrcpprkgaacettrktrrrgcacrlrkclles 128
QY 145 GKKKEMINSDEAVERRALIRKKKSERTGTOPLGVOGLTEORMMIRELMDAOKMTPDIT 204

```



```

Db 129 gmkkemlmsdeaveerrallkxkksertgplvgqilteeqmmirelmdaqrktdt 188
OY 205 FSHKFNRLPGVLSGGCELPESIQAPSRREBAKMSQVAKDCLSKVSLQLRGDSVWNY 264
Db 189 fshfnrlpgvlsaggcelpeslqpsrebaakmsqvdkdclskvslqtrgedgsvwny 248
OY 265 KPPADSGGKEITSLPHMADSTYMKGIISFAKVISYRDLPIEDQISLKGAAFEICQ 324
Db 249 kppadsggkeltslphmadstymkglisfakvisyfrldpliedqislkgafelcq 308
OY 325 LRFNTVNAETGTCGRLSYCLEDTAGGFOQLLEPMKHYMLKKQLHHEEYVLMQA 384
Db 309 lrfntvnaetgwcgrlsyctledtaggfgqlllepmlkhymlkkqlhheeyvlmqa 368
OY 385 ISLSPDRPGVLOHRRVVDLOEOPAITLKSIECNRPQPAHRLFLKIMAMTELRSINA 444
Db 369 islsfdrpgvlyqhrvvdqlgeqfaiitksylecnrppahrfllkimaltelrsina 428
OY 445 OHTQRLRIODIHPRATPLMOELFGITGS 473
Db 429 qhtqrlriodihpratplmqlfgitgs 457

RESULT 7
ID Y25410 standard; Protein; 466 AA.
XX Y25410;
AC Y25410;
DE 06-SEP-1999 (first entry)
XX Human nNR7 partial protein.
DE Human nNR7 partial protein.
XX nNR7; nNR7-1; nuclear trans-acting receptor protein; human; regulator;
XX identification; downstream target gene; cell proliferation;
XX cell development.
XX Homo sapiens.
XX OS
XX PN MO9931129-A1.
XX PD 24-JUN-1999.
XX PF 11-DEC-1998; 98WO-US26364.
XX PR 14-OCT-1998; 98US-0104251.
XX PR 12-DEC-1997; 97US-0069401.
XX PA (MERI) MERCK & CO INC.
XX PI Chen F;
XX WIPI: 1999-405024/34.
XX DR N-PSDB; X78807.
XX PT DNA encoding human nuclear receptors nNR7 and nNR7-1
XX PS Claim 1; Fig 3; 80pp; English.
XX CC This invention describes the isolation of the novel human nuclear
XX receptors nNR7 or nNR7-1. The nNR7 and nNR7-1 proteins are useful in
XX the identification of downstream target genes and ligands regulating its
XX activity. The nuclear receptor is involved in the regulation of in vivo
XX cell proliferation and/or cell development. The nNR7 and nNR7-1
XX polynucleotides, expression vectors and host cells are useful for the
XX recombinant production of the protein.
SO Sequence 466 AA;

```

Query Match 93.6%; Score 2334; DB 20; Length 466;
 Best Local Similarity 98.7%; Pred. No. 3.9e-221;

```

Matches 443; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
OY 25 AAEIASNHPGPPEANTLEVPRKSNHADFVHCETESVPGKPSVNDDEVGQIQIRVCG 84
Db 18 aklpssprgspeanlevrkeshvnhdfhcedtesvpkpsvnddevgqpqirvsg 77
OY 85 DKATYHHNNVTGSCGCGFFPRANKRNARLCPFRKGCETIRKTRROCCARLRKLES 144
Db 78 dkatyhhnnvtcgscgcfprankrnarlcpfrkgcetirkttrroccarlrcles 137
OY 145 GMKKEMINSEAVEERRALTRRKSERTGPIGVGCTEEOXMTRELMDAQMKTFTDT 204
Db 138 gmkkemlnsdeaveerrallkxkksertgplvgqilteeqmmirelmdaqrktdt 197
OY 205 FSHKFNRLPGVLSGGCELPESIQAPSRREBAKMSQVAKDCLSKVSLQLRGDSVWNY 264
Db 198 fshfnrlpgvlsaggcelpeslqpsrebaakmsqvdkdclskvslqtrgedgsvwny 257
OY 265 KPPADSGGKEITSLPHMADSTYMKGIISFAKVISYRDLPIEDQISLKGAAFEICQ 324
Db 258 kppadsggkeltslphmadstymkglisfakvisyfrldpliedqislkgafelcq 317
OY 325 LRFNTVNAETGTCGRLSYCLEDTAGGFOQLLEPMKHYMLKKQLHHEEYVLMQA 384
Db 318 lrfntvnaetgwcgrlsyctledtaggfgqlllepmlkhymlkkqlhheeyvlmqa 377
OY 385 ISLSPDRPGVLOHRRVVDLOEOPAITLKSIECNRPQPAHRLFLKIMAMTELRSINA 444
Db 378 islsfdrpgvlyqhrvvdqlgeqfaiitksylecnrppahrfllkimaltelrsina 437
OY 445 OHTQRLRIODIHPRATPLMOELFGITGS 473
Db 438 qhtqrlriodihpratplmqlfgitgs 466

RESULT 8
ID Y15931 standard; Protein; 434 AA.
XX Y15931;
AC Y15931;
DE 04-AUG-1999 (first entry)
XX A human intranuclear receptor protein.
DE A human intranuclear receptor protein.
XX Human; intranuclear receptor protein; drug development; diagnosis;
XX treatment.
XX Homo sapiens.
XX OS
XX PN JP1127872-A.
XX PD 18-MAY-1999.
XX PF 07-AUG-1998; 98JP-0224172.
XX PR 11-AUG-1997; 97JP-0230335.
XX PA (NIBS) JAPAN TOBACCO INC.
XX PI
XX WIPI: 1999-350330/30.
XX DR N-PSDB; X59966.
XX PT New intranuclear receptor protein - useful for drug development and
XX diagnosis and treatment of disease
XX PS Claim 1; Page 15-16; 38pp; Japanese.
XX CC The present sequence represents a human intranuclear receptor protein.
XX CC library using a swallfish ANO23 derived probe. The protein can
XX CC be used for the development of drugs and diagnosis and treatment
XX CC of various diseases.

```

XX Sequence 434 AA:

Query Match 91.8%; Score 2289; DB 20; Length 434;
Best Local Similarity 100.0%; Pred. No. 9,5e-217;
Matches 434; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

40 LEVPRKESWNADVFHCHDESVGKPSVNADEEVGSPQICRVCGDKATGHHNWTCSG 99
1 LEVPRKESWNADVFHCHDESVGKPSVNADEEVGSPQICRVCGDKATGHHNWTCSG 60
100 CKGFRRAMKNNARLRCFPRGACETIRKTRRQCOACRLKCLSGMKKMMINSDEAVEE 159
61 CKGFRRAMKNNARLRCFPRGACETIRKTRRQCOACRLKCLSGMKKMMINSDEAVEE 120
160 RRALIKRKKSERTGTPGLVGVGLTEEOQMMRELMADQMTFTTFESHFNRLPGVLS 219
121 RRALIKRKKSERTGTPGLVGVGLTEEOQMMRELMADQMTFTTFESHFNRLPGVLS 180
220 GCELPESIQAPSRERAAKMSQVRKDLCSLKYSVLQRGEDGSVMYKPPADSGGKEIFSL 279
181 GCELPESIQAPSRERAAKMSQVRKDLCSLKYSVLQRGEDGSVMYKPPADSGGKEIFSL 240
280 PHMADSTYMPFGIISFAKVISYFRDLPIEDQISILKGAFAELCOLFNTVFNAETGWE 339
241 PHMADSTYMPFGIISFAKVISYFRDLPIEDQISILKGAFAELCOLFNTVFNAETGWE 300
340 CGRLSTCLDEDTAGGFOQLLEPMLKFHYMLKRLQLEEEYVLMQALISLSPDRPGVLOHR 399
301 CGRLSTCLDEDTAGGFOQLLEPMLKFHYMLKRLQLEEEYVLMQALISLSPDRPGVLOHR 360
400 VVDQLOEOPAITLKSYIECNRPQAPARHLEFLKIMMTELINSINAOHTQRLRIODIHPF 459
361 VVDQLOEOPAITLKSYIECNRPQAPARHLEFLKIMMTELINSINAOHTQRLRIODIHPF 420
460 ATPLMOELFGITGS 473
421 ATPLMOELFGITGS 434

RESULT 9
Y09515
ID Y09515 standard; Protein: 437 AA.
XX
AC Y09515;
XX
DT 16-JUL-1999 (first entry)
DE Human vitamin D receptor related gamma protein.
XX
KW Human; vitamin D receptor related protein; VDR; obesity; diabetes;
KW anorexia; rheumatoid arthritis; lipoprotein defect; hyperlipidaemia;
KW hypercholesterolaemia; hyperlipoproteinaemia; osteoporosis; tumour;
KW hyperproliferative skin disorder; hyperthyroidism.
XX
OS Homo sapiens.
XX
PN WO9919354-A1.
XX
PD 22-APR-1999.
XX
PF 31-AUG-1998; 98MO-SF01548.
XX
PR 31-MAR-1998; 98SE-0001148.
PR 14-OCT-1997; 97SE-0003745.
XX
PA (PAAA) PHARMACIA & UPJOHN AB.
XX
PI Berkenstam A, Dahlberg M;
XX
DR WPI; 1999-302508/25.
XX
N-PSDB; X56242.

XX New vitamin D receptor related (VDR) polypeptides, useful for
PT treating obesity, diabetes, anorexia and rheumatoid arthritis
XX
PS Claim 19; Page 19-20; 35pp; English.

CC The present sequence is a human vitamin D receptor related (VDR)
CC polypeptide. Human VDR polypeptides and substances which affect VDR
CC signal transduction, can be used for treating metabolic, proliferative
CC or inflammatory conditions. They can be used in the manufacture of a
CC medicament for treating the following conditions: obesity, diabetes,
CC anorexia, lipoprotein defects, hyperlipidaemia, hypercholesterolaemia or
CC hyperlipoproteinaemia and osteoporosis, rheumatoid arthritis, benign and
CC malign tumours, hyperproliferative skin disorders or hyperthyroidism.
CC Nucleic acid vectors encoding for expression of a VDR polypeptide can
CC be used for treating metabolic, proliferative or inflammatory conditions,
CC by introducing them into a mammal. The introduced nucleic acid is then
CC capable of transforming a cell in vivo and then polypeptide is expressed.
CC A substance affecting VDR signal transduction, such as an agonist or
CC antagonist can be used for the manufacture of a medicament for treating
CC metabolic, proliferative or inflammatory condition.
CC N.B. The specification specifically claims the VDR nucleic acid and
CC polypeptide sequences given in figures 1, 4, 7 and 8, but no figures
XX are given in the specification.

Sequence 437 AA:

Query Match 91.7%; Score 2287; DB 20; Length 437;
Best Local Similarity 99.8%; Pred. No. 1,5e-216;
Matches 433; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

40 LEVPRKESWNADVFHCHDESVGKPSVNADEEVGSPQICRVCGDKATGHHNWTCSG 99
1 LEVPRKESWNADVFHCHDESVGKPSVNADEEVGSPQICRVCGDKATGHHNWTCSG 60
100 CKGFRRAMKNNARLRCFPRGACETIRKTRRQCOACRLKCLSGMKKMMINSDEAVEE 159
61 CKGFRRAMKNNARLRCFPRGACETIRKTRRQCOACRLKCLSGMKKMMINSDEAVEE 120
160 RRALIKRKKSERTGTPGLVGVGLTEEOQMMRELMADQMTFTTFESHFNRLPGVLS 219
121 RRALIKRKKSERTGTPGLVGVGLTEEOQMMRELMADQMTFTTFESHFNRLPGVLS 180
220 GCELPESIQAPSRERAAKMSQVRKDLCSLKYSVLQRGEDGSVMYKPPADSGGKEIFSL 279
181 GCELPESIQAPSRERAAKMSQVRKDLCSLKYSVLQRGEDGSVMYKPPADSGGKEIFSL 240
280 PHMADSTYMPFGIISFAKVISYFRDLPIEDQISILKGAFAELCOLFNTVFNAETGWE 339
241 PHMADSTYMPFGIISFAKVISYFRDLPIEDQISILKGAFAELCOLFNTVFNAETGWE 300
340 CGRLSTCLDEDTAGGFOQLLEPMLKFHYMLKRLQLEEEYVLMQALISLSPDRPGVLOHR 399
301 CGRLSTCLDEDTAGGFOQLLEPMLKFHYMLKRLQLEEEYVLMQALISLSPDRPGVLOHR 360
400 VVDQLOEOPAITLKSYIECNRPQAPARHLEFLKIMMTELINSINAOHTQRLRIODIHPF 459
361 VVDQLOEOPAITLKSYIECNRPQAPARHLEFLKIMMTELINSINAOHTQRLRIODIHPF 420
460 ATPLMOELFGITGS 473
421 ATPLMOELFGITGS 434

RESULT 10
Y21799
ID Y21799 standard; Protein: 434 AA.
XX
AC Y21799;
XX
DT 14-SEP-1999 (first entry)
XX

DE		Human steroid and xenobiotic receptor (SXR).
KW	Nuclear receptor; SXR; steroid and xenobiotic receptor; RXR; human;	
KW	retinoid X receptor; P450 gene; steroid hormone; steroid metabolism;	
KW	phytoestrogen; calcium-channel blocker; steroid toxicity; tuberculosis;	
KW	breast cancer; osteoporosis; Cushing syndrome; virilism; hirsutism;	
KW	polycystic ovarian disease; cancer; colorectal; prostatic.	
OS	Homo sapiens.	
XX		
FH	Key	Location/Qualifiers
FT	Misc-difference 227	/label= unknown
FT		/note= "encoded by ACN"
XX		
PM	W099J35246-A1.	
PD	15-JUL-1999.	
PF	08-JAN-1999; 99WC-US00490.	
PR	09-JAN-1998; 98US-0005286.	
PA	(SALK) SALK INST BIOLOGICAL STUDIES.	
PI	Blumberg B, Evans RM;	
DR	WPI: 1999-419349/35.	
XX	N-PSDB: X89080.	
PT	New steroid and xenobiotic receptor , used to identify modulators for	
PT	controlling metabolism of steroids and xenobiotics, e.g. reducing	
PT	their toxicity	
XX		
PS	Claim 4; Fig 1A; 83pp; English.	
XX		
CC	The invention relates to a novel nuclear receptor polypeptide, designated	
CC	SXR (steroid and xenobiotic receptor). SXR (I) forms a heterodimer with	
CC	retinoid X receptor (RXR), (II) binds to a direct or inverted repeat	
CC	response element motif based on the half-site AGTTCA, (III) activates	
CC	transcription through response elements present in steroid-inducible P450	
CC	genes, in response to a wide variety of natural and synthetic steroid	
CC	hormones and (IV) is prominently expressed in liver and intestine. SXR	
CC	regulates expression of catabolic enzymes, in response to many different	
CC	steroids, and thus affects metabolism. SXR is a broad specificity, low-	
CC	affinity receptor for reducing excessive levels of steroids in the	
CC	circulation. (Anti)agonists of SXR are used to regulate metabolism of	
CC	steroids particularly phytoestrogens or calcium-channel blockers, to	
CC	reduce steroid toxicity in subjects being treated with steroids, e.g. in	
CC	cases of tuberculosis (treated with rifampin and related compounds),	
CC	breast cancer (treated with tamoxifen, raloxifen etc.) or osteoporosis	
CC	(treated with vitamin K), or to slow metabolism of therapeutic steroids.	
CC	Also, modulating endogenous SXR is used to treat disease, particularly	
CC	an agonist is used where endogenous steroid levels are excessive (e.g.	
CC	Cushing syndrome; virilism and hirsutism in women; polycystic ovarian	
CC	disease; II beta-, 17- or 21-hydroxylase deficiency; 3 beta-hydroxysteroidid	
CC	dehydrogenase deficiency, or breast, colorectal or prostatic cancer),	
CC	while antagonists are used where endogenous steroid levels are too low.	
CC	Cells that express SXR are used to identify compounds likely to be	
CC	involved in undesirable drug interactions. Antibodies specific for SXR	
CC	are used in immunochemical testing for studying distribution/	
CC	expression density of SXR, also for diagnosis and therapeutically as	
CC	antagonist. The present sequence represents SXR polypeptide.	
SO	Sequence 434 AA;	

```

Query Match Similarity      87.0%; Score 2171; DB 20; Length 434;
Best Local Similarity      95.9%; Pred. No. 3.8e-205;
Matches 416; Conservative  2; Mismatches 16; Indels 0; Gaps 0.

0y  40 LEVRPESNHNADFYVCECTESYVGRFSVNADEEVGGPQICRCGCGKAGYHHNVATCG 99
|||||

```

```

Db      1 levipkeewr hadfvhcedtesvpgkpsvnaadeevgpgqlcrvgdka tgyhfnvmccg 60
QY      100 CKGFPRRAMKNNALRCPFRKGAEEITRKTRRQCAQLRKLCLESgKKKEMISDAVEE 159
Db      61 ckgffirrmk nna lrcpfrgacelctfrtkrcqcaer l kclcs gmkk emla deavee 120
QY      160 RRALIKRKRSERTSTQPIGVGVLREEDQRMNRLMDLQOMKTFDFTSFHNFRRLPGVLS 219
Db      121 rtal l k r k k e e r t g t p r j g v g l l e e g r m m l r e m d e q m k t f d t f h k n f i p y l s 180
QY      220 GCELPESIQADSRBEAAKWSQVRKDLCSLKVSIQLRGEDGSVWNKYPKPADSGGKEIRSL 279
Db      181 gcelpepqgqsr e e a a k w s q v r k d l c s l k v s l q a e g s g w c l e l x p s r q w k e f s l l 240
QY      280 PHMADMSTNYRKGIISPAKVIYSTRDRLPIEQISLVLGAAPELQALFNFNVAETGWE 339
Db      241 phmdamstnyrk g i i s p a k v i y s t r d r l p i e q i s l v l g a a p e l c q l f n f n v a e t g w e 300
QY      340 CGRLSYCLEETAGGFQOQLLEPMKPFYMLKKTQLDHEEYVLMOAISLSPDRPGVQHR 399
Db      301 cgrlsy c l e e t a g g f q o q l l e p m k p f y m l k k t q l d h e e y v l m o a i s l s p d r p g v q h r 360
QY      400 VVDQLQEQFATTLKSYLECNRPORAHNPLFLKIKAMLTSLRSINAQHTORLRLIODINPF 459
Db      361 vv d q l q e q f a t t l k s y l e c n r p o r a h n p l f l k i k a m l t s l r s i n a q h t o r l r l i o d i n p f 420
QY      460 ATPLMQELFGITGS 473
Db      421 atplmqel f g i t g s 434

```

RESULT 11
Y42691

XX
AC Y42691;

XX
DT 17-JAN-2000 (first entry)
XX

Human pregnane X receptor (hPXR).

KM Human; nuclear receptor; pregnane X receptor; PXR; CYP
 KW cytochrome P-450 mono-oxygenase; drug interaction; hPXR; CYP3A4

xx
os Homo sapiens.

XX
PN W09948915-A1.

XX
PD
30-SEP-1999.

26-MAR-1999, 99WO-US06737.

27-MAR-1998; 98US-0079593.

PA (GLAX) GLAXO GROUP LTD.
XX

PI Kilewer SA, Willson TM;
XX

DR WPI; 1999-601202/51.
DR N-PSDB; Z07997..

XX	New human pregnane X
PT	receptor that induces
PM	

PS C:\atm 4: File 1A-D: 6
XX
FI agencies will induce

XX The invention provides an isolate

CC pregnane X receptor, PXR) that binds to a cytochrome P-450 mono-oxygenase
CC (CYP) promoter. The hPXR is used to identify: its specific modulators,

CC and compounds that induce CYP3A4 expression (i.e. to identify drug
CC interactions, since CYP3A4 is involved in many biotransformations of
CC

CC drugs). The modulators are potentially useful for: associating particular CC diseases and conditions with PXR and for treating such conditions.

CC Antibodies raised against hpxr can be used for determination and
CC purification of hpxr. The present sequence represents the hpxr.

XX Sequence 414 AA;

Query Match 86.9%; Score 2167; DB 20; Length 414;
Best Local Similarity 95.4%; Pred. No. 8.8e-205;
Matches 414; Conservative 0; Mismatches 0; Indels 20; Gaps 1;

QY 40 LEVRKESWNRADVHCEDETSVPKDSVNADEVGGQICRVCGRATGTHNVMTCEG 99
DB 1 Levipkeswnhadfvcdesvypkpsvnaeevgpqlcrvsgdkatgthfvmvceq 60
QY 100 CKGFPRMRNRNARLCGRFRGACETIRKTRCOACOLRCLSGMKKEMINSDEAVEE 159
DB 61 ckgfframkxnarlrccpfrkqaceltrktrcqaccltrcclsgmkkemsdeavee 120
QY 160 RRAIIRKRSRTGTOPFVGILTEQRMIRLMDAQMTFTTFSHFKNFRLPGVLS 219
DB 121 rralikrkksertgtqplvgvgllceeqmmrlmdagmktfdttfshknfrlpgvls 180
QY 220 GCELPESIQADSRBAANWSQVRKDLGSLKVSQLRGDSGVNWKRPADSGKEIRSL 279
DB 181 gcelpesiqadpsrebaakwsgvrkdlcslkvsqllrgedsgsvwnykp padsgkelfsl 240
QY 280 PHMDMSVYMKGIISFPAKVSYFRDLPIEDQISLKGAFELCOLFNTVFNATGTE 339
DB 241 phmdmsvymkglisfakvlsyfrdlpiedqislkgaafelcolfnvfnatgte 300
QY 340 CGRLSYCLEDTAGFQOLLEPMLKFHYMLKKQLHBEYVLMQAISLFSRDRGVQHR 399
DB 301 cgrlsyctledtagfqqllepmkfhymkkqlhbeeyvlmqaislfsrdrpvqhr 360
QY 400 VVDQLOQFOAITLKSITCNRPQPAHRFLFKIMAMTELRSINAQTHORLRIQDIHPF 459
DB 361 vvdqlqefaitlksylecnrpqahrflfkimamtelrsinaqthorlriqdihp f 400
QY 460 ATPLMQELFGITGS 473
DB 401 atplmqelfgits 414

RESULT 12

Y42689
ID Y42689 standard; Protein; 316 AA.

XX Y42689;

DT 17-JAN-2000 (first entry)

DE His6-pregnane X receptor (PXR) sequence.

KW Human; nuclear receptor; pregnane X receptor; PXR; CYP; CYP3A4;
KW cytochrome P-450 mono-oxygenase; drug interaction; hpxr.

OS Synthetic.

XX Homo sapiens.

XX W09948915-A1.

PD 30-SEP-1999.

PE 26-MAR-1999; 99WO-US06737.

FR 27-MAR-1998; 98US-0079593.

PA (GLAX) GLAXO GROUP LTD.

PI Kilewer SA, Willson TM;

XX WPI; 1999-601202/51.

PI New human pregnane X receptor, used to identify specific modulators and
PI agents that induce expression of cytochrome P-450 mono-oxygenase
XX Example 6; Page 37; 69pp; English.

CC The invention provides an isolated human nuclear receptor (designated
CC pregnane X receptor, PXR) that binds to a cytochrome P-450 mono-oxygenase
CC (CYP) promoter. The hpxr is used to identify its specific modulators,
CC and compounds that induce CYP3A4 expression (i.e. to identify drug
CC interactions, since CYP3A4 is involved in many biotransformations of
CC drugs). The modulators are potentially useful for associating particular
CC diseases and conditions with PXR and for treating such conditions.
CC Antibodies raised against hpxr can be used for determining and
CC purification of hpxr. The present sequence represents a histidine-6
CC tagged partial PXR (His6-PXR) sequence.

XX Sequence 316 AA;

Query Match 63.6%; Score 1585; DB 20; Length 316;
Best Local Similarity 100.0%; Pred. No. 1e-147;
Matches 305; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 169 SERTGTQPLGVOGLTEQRMIRLMDAQMTFTTFSHFKNFRLPGVLSGCELPESLQ 228
DB 12 sertgtqplvgvgllceeqmmrlmdagmktfdttfshknfrlpgvlsqcelpeslq 71
QY 229 APSREBAANWSQVRKDLGSLKVSQLRGDSGVNWKRPADSGKEIRSLPHMDMSVY 288
DB 72 apsebaanwsqvrkdlcslkvsqllrgedsgsvwnykp padsgkelfslphmdmsvly 131
QY 289 MFKGISFPAKVSYFRDLPIEDQISLKGAFELCOLFNTVFNATGTEWCGRLSYCLE 348
DB 132 mfkglisfakvlsyfrdlpiedqislkgaafelcolfnvfnatgtecgrlsycle 191
QY 349 DTAGSQOLLEPMLKFHYMLKKQLHBEYVLMQAISLFSRDRGVQHRVVDQLOQF 408
DB 192 dtagfqqllepmkfhymkkqlhbeeyvlmqaislfsrdrpvqhrvvdqlqef 251
QY 409 AITLKSYICNRPQPAHRFLFKIMAMTELRSINAQTHORLRIQDIHPATPLMQELF 468
DB 252 aitlksylecnrpqahrflfkimamtelrsinaqthorlriqdihp atplmqelf 311
QY 469 GITGS 473
DB 312 gitgs 316

RESULT 13

R98521
ID R98521 standard; Protein; 386 AA.

AC R98521;

DT 14-NOV-1996 (first entry)

DE Xenopus orphan receptor 6.

KW Xenopus orphan receptor 6; XOR-6; steroid receptor; vitamin D;
KW hydroxybenzoate; mercaptohydroxybenzoate; aminobenzoate.

XX Xenopus laevis.

OS Xenopus laevis.

XX Key

FT Domain

FT Domain

FT Domain

PN W09622390-A1.

XX 25-JUL-1996.

Location/Qualifiers
/Label= DNA_Binding_domain
/Label= 183..386
/Label= Ligand_Binding_domain

PF 16-JAN-1996; 96WO-US00058.
 XX
 PR 17-JAN-1995; 95US-0374445.
 XX
 XX (SALK) SALK INST BIOLOGICAL STUDIES.
 PA
 XX Blumberg B, Evans RM, Umehono K;
 PI
 XX WPI: 1996-354546/35.
 DR
 XX N-PSDB; T36499.
 XX
 PT DNA encoding receptor polypeptide responsive to hydroxy, mercapto or
 amino benzothiazole(s) - useful to regulate gene transcription
 PS
 PS Claim 4; Page 27-28; 42pp; English.
 XX
 CC Xenopus orphan receptor 6 (XOR-6) (R98521) is a new member of the
 CC steroid receptor superfamily, characterized as being responsive to
 CC the presence of hydroxy, mercapto or amino benzothiazole(s) and as
 CC regulating the transcription of associated gene(s). It shows 73%
 CC identity in the DNA binding domain, and 42% identity in the ligand
 CC binding domain, to the human vitamin D receptor. Recombinant XOR-6
 CC can be expressed in animal cells; a cDNA clone (T36499) coding for
 CC XOR-6 has been isolated. The recombinant XOR-6 may be used to
 CC regulate gene transcription or to raise antibodies of diagnostic
 CC or therapeutic appln.
 CC
 XX
 SQ Sequence 386 AA;

Query Match 39.3%; Score 979.5; DB 17; Length 386;
 Best Local Similarity 48.6%; Pred. No. 5.1e-88;
 Matches 202; Conservative 60; Mismatches 105; Indels 49; Gaps 9;

YY 57 EDTESVPGKPSVANDEVEGPOICRVGCDKATGYHFNVTCEGKGFRRAMKRNALNC 116
 DB 14 eeedaenscgtdedgdgpkicrcgdratgyhfnantcegcgkffrravrnrlisc 73
 YY 117 PFRKACETIRKTRKOCACRLKCLSEGMKMKMSDEAVEERALLKRR-KSERGTGQ 175
 DB 74 pf-qnscevlnknnrrncgeerlkkcdlgnrkclmsdaavegrallkrrhkltppt 132
 YY 176 PLGVGGLTEORRMRELMDAOKTDTFTHKFNRLPGVLSGCELPESLOAPRENA 235
 DB 133 ppea-qlpeqgqfllqvgahktldfnfctskntr-----plr--- 171
 YY 236 AKWSQVRKDLCSLKSLSQLRGEGSVWYKPPADSGKKEIFSLPHMADSTYMKGIS 295
 DB 172 -----rsadpt---qepqets--seafimlphladlvygmkgilis 207
 YY 296 FAKVISYFRDLPEDQISLKGAFELCOLRFTVFNAGETGWEGRSLSYCLEDT-AGGF 354
 DB 208 fammlyfslsldedqiallkgsvaeavylrftvfnadctwecgpflydcdmllagf 267
 YY 355 QOQLLEPMKLFHMLKRLQLEHEEYVLMQALISLSPDRPGVLOHRYVDOLOPAILTKS 414
 DB 268 rqlflepvlrthmmklnvgeeyammaaiaifaayrpgvcdwekqlqghlailclxd 327
 YY 415 YIECAN-PPAHREFLKTKMAMTELRSINAQHTORLRIODIHPPARLMOELRG 469
 DB 328 fidsqppspqprillypkimeclteitrvndhskqllelwdlqgdaqlmzevg 383

RESULT 14
 W94623
 ID W94623 standard; Protein; 423 AA.
 XX
 AC W94623;
 XX
 XX 29-APR-1999 (first entry)
 XX
 DE Rat vitamin D receptor protein VDR0.
 XX

KM Vitamin D receptor; VDR: bone density; retinoic acid derivative;
 KM steroid; bone calcium regulator; immunosuppressant; anticancer.
 XX
 XX Rattus sp.
 XX
 XX WO9856908-A1.
 XX
 XX
 PD 17-DEC-1998.
 XX
 XX 13-JUN-1997; .97WO-JP02052.
 PF
 XX 13-JUN-1997; 97WO-JP02052.
 PR
 XX 13-JUN-1997; 97WO-JP02052.
 XX
 PA (CHUS) CHUGAI SEIYAKU KK.
 XX
 XX Kato S, Ueno K;
 PI
 XX WPI: 1999-080898/07.
 DR
 XX N-PSDB; X16606.
 DR
 XX
 PT Gene for vitamin D receptor isoform protein which blocks vitamin D
 PT signal pathway - and production of recombinant isoform protein using
 PT it, for bone density assay and for screening compounds for vitamin D
 PT activity.
 PS
 PS Example 1; Fig 1; 47pp; Japanese.
 XX
 CC The present invention describes rat vitamin D receptor (VDR) isoform
 CC protein VDR1. The present sequence represents rat VDR0. VDR1 differs
 CC from the canonical form of VDR (VDR0) by lacking the amino acid sequence
 CC encoded by exon 9, but instead substitutes a short sequence
 CC (GTEPGSEELRDGHVGDCE in the rat protein) encoded by the 5'-end of intron
 CC 8. VDR1 has a dominant negative effect on the vitamin D signalling
 CC pathway. The isoform protein can be used for determining bone density
 CC and for the screening of compounds (e.g. steroids and retinoic acid
 CC derivatives) for vitamin D activity (e.g. as bone calcium regulators,
 CC immunosuppressants or anticancer agents).
 CC
 XX
 SQ Sequence 423 AA;

Query Match 32.1%; Score 801; DB 20; Length 423;
 Best Local Similarity 42.0%; Pred. No. 2.1e-70;
 Matches 172; Conservative 75; Mismatches 135; Indels 28; Gaps 9;

YY 77 POICRVGCDKATGYHFNVTCEGKGFRRAMKRNALRCRPFKACETIRKTRKOCAC 136
 DB 21 plicgvcgdratgyhfnantcegcgkffrrsmkralfcpf-ngdrrlkdrrhcgac 79
 YY 137 RLKRLCSGMKMKMSDEAVEERALLKRRKSRRTGTOPLGVOGLTEORRMRELMDA 196
 DB 80 rlcrcvdlgmkkellldeevqrrkrrimrkceeaakdrlrp-lseegqnlallda 138
 YY 197 QMKTFTDTSFGKFNRLPGV--SSGCELPESLOAPRENAKWSQVRKDLCSLV--- 250
 DB 139 hbktydptlyadlfrpvrmdsgtsyspr---plsfsgnssassadlytsldme 194
 YY 251 -----SLOLGEDESVWYKPPADSGKKEIFSLPHMADSTYMKGISFAVISYFRD 305
 DB 195 psgfnsldlgedd-----dpsvclldlpslmlphladlvsyqlkvlgfakm1pgfid 249
 YY 306 LPEDQISLKGAFELCOLRFTVFNAGETGWEGRSLSYCLEDT--AGGFQOQLLEPM 362
 DB 250 ltsddqivllksaalevimlrnsqsfmdmdmawdcgsqdykyvdtvskaghlellep 309
 YY 363 LKHFHMLKRLQLEHEEYVLMQALISLSPDRPGVLOHRYVDOLOPAILTKSIECNRPQ 422
 DB 310 lkrtvgllklnlneeevllmalclvpspdrpyvdaklvealqdrlnslqlyircrhp 369
 YY 423 PAHREFLKTKMAMTELRSINAQHTO--RLLRIDQIHPP-APRLMOELRG 469
 DB 370 pgaqlyakmlqladlrlnseehakysrlsfqpsnsmklplvlevrg 419

```

RESULT 15
W47509
ID W47509 standard; Protein: 423 AA.
XX
AC W47509;
XX
DE 26-JUN-1998 (first entry)
XX
DE Rat vitamin D receptor (VDR).
XX
XX Rat; vitamin D receptor; isoform protein; VDR1; VDR0; diagnosis;
XX dominant negative receptor; signal transmission channel;
XX bone density disorder; screening.
XX
OS Rattus rattus.
XX
FH Key Location/Qualifiers
FT Misc-difference 360
FT /note="encoded by CAG"
XX
XX MO9747172-A1.
XX
PD 18-DEC-1997.
XX
XX 10-JUN-1997; 97WO-IB00947.
XX
XX 10-JUN-1996; 96JP-0194179.
XX
XX (CHUS ) CHUGAI PHARM CO LTD.
XX (CHUS ) CHUGAI SEIYAKU KK.
XX
XX Kato S, Ueno K;
XX
XX WPI: 1998-051917/05.
XX N-PSDB; V03129.
XX
XX DNA encoding a vitamin D receptor isoform protein - useful for bone
XX density determination and for screening substances for vitamin D
XX activity
XX
XX PS Disclosure; Fig 1; 46pp; Japanese.
XX
XX A novel cDNA sequence encodes the rat vitamin D receptor isoform
XX protein (VDR1). The isoform differs from the normal receptor
XX (VDR0), which comprises the present sequence, in having the
XX vitamin D response element curtailed by 86 residues, and having an
XX extra 19 residues inserted at the C-terminal of this element. It
XX acts as a dominant negative receptor in the vitamin D signal
XX transmission channel.
XX The isoform protein can be used to diagnose bone density disorders,
XX and screen for substances having potential vitamin D-like activity.
XX
XX Sequence 423 AA;

```

```

Query Match 32.0%; Score 798; DB 19; Length 423;
Best Local Similarity 42.0%; Pred. No. 4,2e-70;
Matches 172; Conservative 74; Mismatches 136; Indels 28; Gaps 9;

```

```

OY 77 POICVCGDKATGYHFNVTCEGCGEPPRRAMRNALRCPFRKACETTRTRRQOAC 136
DB 21 prlgvcgdratgfnamtecgckgffirmkrltfcpf-ngdcritkdnrrhcqac 79
OY 137 RLKRCLESGMKEMIMSDAVERALIKKKSEKTOPGLGVOGLTFEORMTREMDA 196
DB 80 rlkrcvldgmkefllideevqkrkemlmkreeealldslrpx-lseegqhlallda 138
OY 197 QMKTEDTFESHKRNRLGVL--SSGCELPESLQAPSRREAKMSQVRKDCSLK-- 250
DB 139 hnktydpyadrtidrrpymdsgyspr-----plsfsgnsasssdltytsldme 194
OY 251 -----SLQIRGEDGSVMWKKPPADSGKEIFSLPLPHMADMTYMFKGIISFAKVI 305

```

```

DB 195 psgfndldngedsd-----dpavtldlplsmplphladivsyslqkylgafakmipgrfd 249
OY 306 LPREDQISLKGAAFEICQURFNTVNAETGTEGCRSLSCLEDT--AGGFOQLLEPM 362
DB 250 ltsdqqivllkssalevmlrnsqsfmdmswcdsgsdykxvdtavskaghtlelepl 309
OY 363 LKPHYMLKTLQHEEYVLMQATSLFSPDRPGYLQHRVVDLOEQPAILTKSYIECNRQ 422
DB 310 krlgvgllklnlneehvllmalcivspdrpgvdaklveaigdrisnlgtylrcrhp 369
OY 423 PAHRFLFKTAMLTLELSINAQHTQ--RLRLIQDHPF-ATPLMQELRG 469
DB 370 psgnqlyakmqladlrslneehskyzslsfqpsnmkltpdlvlevlg 419

```

```

Search completed: February 18, 2001, 14:29:40
Job time: 8217 sec

```

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 18, 2001, 14:35:36 ; Search time 3075.67 Seconds
(without alignments)
6618.640 Million cell updates/sec

Title: US-09-143-828-1

Perfect score: 2905
Sequence: 1 cctctgaaggctctagaatc.....aaaaaaaaaaaaaaaaaaaaa 2905

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 7991742 seqs, 3503743858 residues

Total number of hits satisfying chosen parameters: 15983484

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST.*

1: gb_est1.*
2: gb_est2.*
3: gb_est3.*
4: gb_est4.*
5: gb_est5.*
6: gb_est6.*
7: gb_est7.*
8: gb_est8.*
9: gb_est9.*
10: gb_est10.*
11: gb_est11.*
12: gb_est12.*
13: gb_est13.*
14: gb_est14.*
15: gb_est15.*
16: gb_est16.*
17: gb_est17.*
18: gb_est18.*
19: gb_est19.*
20: gb_est20.*
21: gb_est21.*
22: gb_est22.*
23: gb_est23.*
24: gb_est24.*
25: gb_est25.*
26: gb_est26.*
27: gb_est27.*
28: gb_est28.*
29: gb_est29.*
30: gb_est30.*
31: gb_est31.*
32: gb_est32.*
33: gb_est33.*
34: gb_est34.*
35: gb_est35.*
36: gb_est36.*
37: gb_est37.*
38: gb_est38.*
39: gb_est39.*
40: gb_est40.*
41: em_estba.*
42: em_estun1.*
43: em_esthum1.*

44: em_esthum2.*
45: em_esthum3.*
46: em_esthum4.*
47: em_esthum5.*
48: em_esthum6.*
49: em_esthum7.*
50: em_esthum8.*
51: em_esthum9.*
52: em_esthum10.*
53: em_esthum11.*
54: em_esthum12.*
55: em_esthum13.*
56: em_esthum14.*
57: em_esthum15.*
58: em_esthum16.*
59: em_esthum17.*
60: em_esthum18.*
61: em_esthum19.*
62: em_esthum20.*
63: em_est101.*
64: em_est102.*
65: em_est103.*
66: em_est104.*
67: em_est105.*
68: em_est106.*
69: em_est107.*
70: em_est108.*
71: em_est109.*
72: em_est110.*
73: em_est111.*
74: em_est112.*
75: em_est113.*
76: em_est114.*
77: em_est115.*
78: em_est116.*
79: em_est117.*
80: em_est118.*
81: em_est119.*
82: em_est120.*
83: em_est121.*
84: em_est122.*
85: em_est123.*
86: em_est124.*
87: em_est125.*
88: em_est126.*
89: em_est127.*
90: em_est128.*
91: em_est129.*
92: em_est130.*
93: em_est131.*
94: em_est132.*
95: em_est133.*
96: em_est134.*
97: em_est135.*
98: em_est136.*
99: em_est137.*
100: em_est138.*
101: em_est139.*
102: em_est140.*
103: em_est141.*
104: em_est142.*
105: em_est143.*
106: em_est144.*
107: em_est145.*
108: em_est146.*
109: em_est147.*
110: em_est148.*
111: em_est149.*
112: em_est150.*
113: em_est151.*
114: em_est152.*
115: em_est153.*
116: em_est154.*
117: em_est155.*
118: em_est156.*
119: em_est157.*
120: em_est158.*
121: em_est159.*
122: em_est160.*
123: em_est161.*
124: em_est162.*
125: em_est163.*
126: em_est164.*
127: em_est165.*
128: em_est166.*
129: em_est167.*
130: em_est168.*
131: em_est169.*
132: em_est170.*
133: em_est171.*
134: em_est172.*
135: em_est173.*
136: em_est174.*
137: em_est175.*
138: em_est176.*
139: em_est177.*
140: em_est178.*
141: em_est179.*
142: em_est180.*
143: em_est181.*
144: em_est182.*
145: em_est183.*
146: em_est184.*
147: em_est185.*
148: em_est186.*
149: em_est187.*
150: em_est188.*
151: em_est189.*
152: em_est190.*
153: em_est191.*
154: em_est192.*
155: em_est193.*
156: em_est194.*
157: em_est195.*
158: em_est196.*
159: em_est197.*
160: em_est198.*
161: em_est199.*
162: em_est200.*

117: em_estp16:*
 118: em_estp17:*
 119: em_estp18:*
 120: em_estp19:*
 121: em_estp20:*
 122: em_estp21:*
 123: em_estp22:*
 124: em_estp23:*
 125: em_estp24:*
 126: em_estp25:*
 127: em_estp26:*
 128: em_estp27:*
 129: em_estp28:*
 130: em_estp29:*
 131: em_estp30:*
 132: em_estp31:*
 133: em_estp32:*
 134: em_estp33:*
 135: em_estp34:*
 136: em_estp35:*
 137: em_estp36:*
 138: em_estp37:*
 139: em_estp38:*
 140: em_estp39:*
 141: em_estp40:*
 142: em_estp41:*
 143: em_estp42:*
 144: em_estp43:*
 145: em_estp44:*
 146: em_estp45:*
 147: em_estp46:*
 148: em_estp47:*
 149: em_estp48:*
 150: em_estp49:*
 151: em_estp50:*
 152: em_estp51:*
 153: em_estp52:*
 154: em_estp53:*
 155: em_estp54:*
 156: em_estp55:*
 157: em_estp56:*
 158: em_estp57:*
 159: em_estp58:*
 160: em_estp59:*
 161: em_estp60:*
 162: em_estp61:*
 163: em_estp62:*
 164: em_estp63:*
 165: em_estp64:*
 166: em_estp65:*
 167: em_estp66:*
 168: em_estp67:*
 169: em_estp68:*
 170: em_estp69:*
 171: em_estp70:*
 172: em_estp71:*
 173: em_estp72:*
 174: em_estp73:*
 175: em_estp74:*
 176: em_estp75:*
 177: em_estp76:*
 178: em_estp77:*
 179: em_estp78:*
 180: em_estp79:*
 181: em_estp80:*
 182: em_estp81:*
 183: em_estp82:*
 184: em_estp83:*
 185: em_estp84:*
 186: em_estp85:*
 187: em_estp86:*
 188: em_estp87:*
 189: em_estp88:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	523.2	18.0	638	38	AV651714
2	445.2	15.3	461	10	AA699679
3	416	14.3	454	13	N94132
4	370.8	12.8	794	24	A1746915
5	343	11.8	343	18	A1248626
6	335	11.5	365	143	N69311
7	201	6.9	472	4	AA277370
8	199.6	6.9	493	96	AM871811
9	185.4	6.4	200	10	AA679591
10	166.4	5.7	463	91	AM511148
11	165.4	5.7	215	31	AV108557
12	156.8	5.4	432	40	AM158294
13	154.4	5.3	601	40	AM107536
14	143.8	5.0	560	24	A1768052
15	140.8	4.8	1053	110	BE573656
16	132.8	4.6	540	38	AM012320
17	132	4.5	463	15	A1049299
18	118	4.1	724	15	A1046425
19	108.4	3.7	468	6	AA396982
20	102	3.5	463	2	AA107961
21	101.2	3.5	506	10	AA679391
22	100.8	3.5	616	17	A1238397
23	99.2	3.4	732	28	AJ393915
24	99	3.4	370	15	A1024672
25	99	3.4	507	4	AA264990
26	97.6	3.4	690	106	BE266478
27	97.6	3.4	1105	107	BE410116
28	97.2	3.3	589	106	BE275249
29	94	3.2	851	137	BE877201
30	90.2	3.1	708	107	BE390042
31	88.6	3.1	692	135	BE791424
32	88.6	3.0	392	91	AM505438
33	88.6	3.0	475	111	BE664197
34	86.6	3.0	539	16	A1098643
35	83.6	2.9	623	106	BE303050
36	83.4	2.9	788	137	BE915476
37	82.8	2.9	665	107	BE390515
38	82.4	2.8	467	9	AA570839
39	81.8	2.8	959	137	BE903117
40	81.4	2.8	564	37	AV609965
41	81.2	2.8	771	107	BE385638
42	81	2.8	248	4	AA272757
43	81	2.8	514	2	AA124058
44	80.4	2.8	667	87	AM231123
45	80	2.8	563	105	BE233303

ALIGNMENTS

RESULT 1
 AV651714
 LOCUS AV651714 638 bp mRNA
 DEFINITION AV651714 GIC Homo sapiens CDNA clone GICCS03 3', mRNA sequence.
 ACCESSION AV651714
 VERSION AV651714.1 GI:9872728
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens.

This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
MGI:994718

Possible reversed clone: similarity on wrong strand
Seq primer: custom primer used
High quality sequence stop: 472.

FEATURES

source

1..794
Location/Qualifiers
/organism="Mus musculus"
/strain="C57BL"
/db_xref="taxon:10090"
/clone="IMAGE:2076282"
/clone_1lb="Sugano mouse embryo mewa"
/dev_stage="embryo, 14 dpc"
/lab_host="DH10B"
/note="Vector: pME18S-FL3; Site 1: DraIII (CACTGTGTC);
Site 2: DraIII (CACCATGTC); 1st strand cDNA was primed
with an oligo(dT) primer [ATGTGGCTTTTCTTTTCTTTT];
double-stranded cDNA was ligated to a DraIII adaptor
[TGTGGCTTCTG], digested and cloned into distinct DraIII
sites of the pME18S-FL3 vector (5' site CACTGTGTC, 3' site
CACCATGTC). XhoI should be used to isolate the cDNA
insert. Size selection was performed to exclude fragments
1.5kb. Library constructed by Dr. Sumio Sugano
(University of Tokyo Institute of Medical Science).
Custom primers for sequencing: 5' end primer
CTTCTCTTAAAGCTGCG and 3' end primer
CGACCTGCAGCTCAGACACA."
BASE COUNT 206 a 180 c 253 g 152 t 3 others
ORIGIN

Query Match

Best Local Similarity 12.8%; Score 370.8; DB 24; Length 794;
Matches 451; Conservative 0; Mismatches 115; Indels 3; Gaps 1;

QY 179 ctgtatcgatccttgcacggatgttcaaaagtggaccccgagggagagtcgagacaa 238
DB 229 CTGGCGACGAACATCTCTCTCTCTCAAGTGGAGCCCAAGGGGAAATCCAAACAA 288
QY 239 agacttaccacacagcaatccaaagggcccgaaagcaacttgagagtgaacccaag 298
DB 289 AGAGTGGCCCCAGACACTTGAACACACAGATTAACCTAAGATGAGAACCTCGAG 348
QY 299 aaagctggaacacatgtctgtacactgtgagagacacagagctgttctctgaaagc 358
DB 349 AGAGCTGAGAGCGAGTGGCTTGTACATGTGAAGAACACACCTCTGCGAAGAGC 408
QY 359 ccaggttcaacgagatgaggaagtcggaggtcccaaatctgcgtgtatgtgggaca 418
DB 409 CCA--TCAACGTAGAGGAGGAAGATGAGGTCTTCAATCTGCGTGTATGTGGGACA 465
QY 419 agggccactggctacatcctcaatgtcattgacatgtgaagatgacaaaggtctttcaga 478
DB 466 AGGCCAATGGCTTACCACTTCAATGTATAGCTGTGAAGAGCAAGGGGTTTTTCAGAA 525
QY 479 gggccatgaaacgcaacgcccggctgaggtgccccttccggaagggcgccctgagatca 538
DB 526 GGGCCATGAAGCATGTCCGGCTGAGGTGCCCTTCGCAAGGAACCTCGAGATCA 585
QY 539 ccgggaagacccggcgagatgcagagcttgcgcctcgcgcaagtgtcctggaagcgga 598
DB 586 CCCGGAAGACAGACGCGAGTCCAGAGCTCGCGTTTCCGCAAGTCCCGAGAGATGACA 645
QY 599 tgaagaagaagatgatacatgtccagagagcggtggaagagggggggtcttgatcagac 658
DB 646 TGAAGAAAGATGATCATGATCCGATCCGCTGTGAGAGAGAGGGGGCTTGATCAAGA 705
QY 659 ggaagaagaatgaaacgagagctcagccacatgaggaatgcaagggctgacagagggac 718
DB 706 GGAAGAAAGAGGAANAGATTGAGGCTCCACGCGCTGGGAGGAGGAGGAGGAGGAGAAC 765
QY 719 agcgatgatgatcaggaagctgatatgac 747

DB 766 AGCAAGCGCTGATCCAGAGCTGATGAC 794

RESULT 5
AI248626/c 343 bp mRNA EST 01-DEC-1998
LOCUS qh77b12.x1 Soares fetal liver spleen INFLS.S1 Homo sapiens cDNA
DEFINITION clone IMAGE:1850687 3' similar to gb:003258 VITAMIN D3 RECEPTOR
(HUMAN); mRNA sequence.

ACCESSION AI248626

VERSION AI248626.1 GI:3844023

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 343)

NCI-CCG http://www.nci.nih.gov/ncicgap.

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

Unpublished (1997)

Contact: Robert Strausberg, Ph.D.

Tel: (301) 496-1550

Email: Robert.Strausberg@nih.gov

This clone is available royalty-free through LNL; contact the

IMAGE Consortium (info@image.lnl.gov) for further information.

Insert Length: 608 Std Error: 0.00

Seq primer: -400P from Glibco

High quality sequence stop: 338.

FEATURES

source

1..343
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1850687"
/clone_1lb="Soares fetal liver spleen INFLS.S1"
/sex="male"
/dev_stage="20 week post conception fetus"
/lab_host="DH10B (empicillin resistant)"
/note="Organ: Liver and Spleen; Vector: pT73D (Pharmacia)
with a modified polylinker; Site 1: Pac I; Site 2: Eco RI;
This is a subtracted version of the original Soares fetal
liver spleen INFLS library. 1st strand cDNA was primed
with a Pac I - oligo(dT) primer [5'
AAGTGAAGATTAATTAAGATCTTTTCTTTTCTTTTCTTTT 3'].
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Pac I and cloned into the Pac I
and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization. Library
constructed by Bento Soares and M.Fatima Bonaldo."
BASE COUNT 56 a 108 c 93 g 86 t
ORIGIN

Query Match

Best Local Similarity 11.8%; Score 343; DB 18; Length 343;
Matches 343; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 255 cagtcgaagggcccgagacaaacttgagagtgagaccgaagaagctggaacatgc 314
DB 343 CAGTCCAAAGGGCCCAAGAACCAACCTGGAGTGAGACCCCAAGAAAGCTGGAACATGC 284
QY 315 tgacttgtaactgtgtagagacacagaggtgttctctggaagcccaagtgtaaacgaga 374
DB 283 TGAATTGTACCTGTGAGGACACAGAGTGTCTCTGGAAGCCCAAGTCTCAACCAAGA 224
QY 375 tgaagaagtgagaggttcccaaaacttgcgtgtatgtgagggaagggccactgcatca 434
DB 223 TGAAGAGTGGAGGTGCCCAAAATCTGCCGTGTATGTGGGAGAACAGCCACTGCTATCA 164
QY 435 cttaaatgatacatgataatgaaagtgcaagggcttttcaagagggccaatgaacgaa 494
DB 163 CTTCATATGTCATGACATGTGAAGATGCAAGGCGCTTTTTCAGAGAGGGCCATGAACGCAA 104

QY 495 cccccgctgtagtggtgcccccttcggaagggcgctgtagatcacccggaagaccggcg 554
 Db 103 CGCCCGGTGAGGTGCCCCCTTCCTCCGAAAGGGCGCTGCGAGATCACCCGGAAGACCCGGCG 44
 QY 555 aacatgcacagcctgcgcgcctgcgaagtgcttgagagcg 597
 Db 43 ACACTGCCAGGCCCTGCGCCCTGCGCAAGTGCCTGAGAGCGGC 1

RESULT 6
 LOCUS N69311 365 bp mRNA EST 13-MAR-1996
 DEFINITION ra2506.s1 Soares fetal liver spleen INFLS Homo sapiens cDNA clone
 IMAGE:293626 3', mRNA sequence.
 ACCESSION N69311
 VERSION N69311.1 GI:1225472
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 365)
 AUTHORS Haller, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M., Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F., Trevasakis, E., Waterston, R., Williamson, A., Wohlmann, P. and Wilson, R.
 TITLE The Washu-Merc EST Project
 JOURNAL Unpublished (1995)
 COMMENT Contact: Wilson RK
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: estewatson.wustl.edu
 This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
 Seq primer: ml3 -40 forward
 High quality sequence stop: 279.

FEATURES
 SOURCE Location/Qualifiers
 1..365
 /organism="Homo sapiens"
 /db_xref="GDB:3801473"
 /db_xref="taxon:9606"
 /clone="IMAGE:293626"
 /clone_lib="Soares fetal liver spleen INFLS"
 /sex="male"
 /dev_stage="20 week-post conception fetus"
 /lab_host="DH10B (ampicillin resistant)"
 /note="Organ: Liver and Spleen; Vector: pRT73D (Pharmacia) with a modified polylinker; Site 1: Pac I; Site 2: Eco RI; 1st strand cDNA was primed with a Pac I - oligo(dT) primer [5' ACTCGAAGAAATTAATTAAGACTTTTCTTTTCTTTTCTTTT 3'], double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Pac I and cloned into the Pac I and Eco RI sites of the modified pRT73 vector. Library went through one round of normalization. Library constructed by Bento Soares and M. Fatima Bernaldo."

BASE COUNT 95 a 106 c 68 g 95 t 1 others
 ORIGIN

Query Match 11.5%; Score 335; DB 143; Length 365;
 Best Local Similarity 97.0%; Pred. No. 2.5e-67;
 Matches 352; Conservative 0; Mismatches 10; Indels 1; Gaps 1;
 QY 2354 aaggaagaaatgataatgacaaaagc-agcacaaggaattccctgtgtgtagctgag 2412
 Db 365 AAAGGAGAAATGATTAAGTACAAAGCAACCAAGAAATTTCCCTGTGTGATGCTGAG 306
 QY 2413 ctgtgtgtgtagcagcagctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2472
 Db 305 CTGTGATGTGGCGGCACTGCGGTGTAACCAAGTGAAGTTCCTCCGAGGACATGATGCTGTAGAG 246

QY 2473 caaggacacacactgcaactgtgtagtgctgtgtgtgtgtgtgtgtgtgtgtgtgt 2532
 Db 245 CAAGGGCAACAACTGACACTGTGAGTGGCTGTGTGTGTGTGTGTGTGTGTGTGT 186
 QY 2533 gccactgtatggggcctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2552
 Db 185 GCACTGTATGGGGCTGGGCTTGTCTGTGGGGCTGGAATCTGGGTATGCTGTGTGACG 126
 QY 2593 aggtcagctgtgcaatcagtttaaacacccggaagaaacatttaacagcttatat 2652
 Db 125 AAGGTACGCTTACAAATCAAGTTAAACACCGAGAGAACATTTACATGCACTTATAT 66
 QY 2653 tctgtgtacacatctattctcaagcctaaaggtatgaagtgcctgtgtgtgtgtgt 2712
 Db 65 TTCTGTGTACATCTATTCTCAAAAGCTAAAGGTATGAAGTGCCTGCTTTGTATAG 6
 QY 2713 cca 2715
 Db 5 CCA 3

RESULT 7
 LOCUS AA277370 472 bp mRNA EST 01-APR-1997
 DEFINITION va81a12.f1 Soares mouse NMU Mus musculus cDNA clone IMAGE:73758 5' similar to TR:G410518 G410518 ORPHAN NUCLEAR RECEPTOR OF STEROID/THYROID SUPERFAMILY. ; mRNA sequence.
 ACCESSION AA277370
 VERSION AA277370.1 GI:1917850
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murine; Mus.

REFERENCE 1 (bases 1 to 472)
 AUTHORS Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T., Gesell, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M., Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B., Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and Waterston, R.
 TITLE The Washu-HMI Mouse EST Project
 JOURNAL Unpublished (1996)
 COMMENT Contact: Marra M/Mouse EST Project
 Washu-HMI Mouse EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@watson.wustl.edu
 This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:454806
 Putative full length read
 vector to vector length is 506
 Seq primer: -28ml3 rev2 Et from Amersham.

FEATURES
 source Location/Qualifiers
 1..472
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /clone="IMAGE:73758"
 /clone_lib="Soares mouse NMU"
 /tissue_type="Liver"
 /lab_host="DH10B"
 /note="Vector: pRT73D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTCACATCTGAGTGGAGCGCGCGCGAATCTTTTCTTTTCTTTT 3']; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pRT73 vector. Library constructed and normalized by Bento Soares and M. Fatima Bernaldo."

QY 2473 caaggacacacactgcaactgtgtagtgctgtgtgtgtgtgtgtgtgtgtgtgtgt 2532
 Db 245 CAAGGGCAACAACTGACACTGTGAGTGGCTGTGTGTGTGTGTGTGTGTGTGTGT 186
 QY 2533 gccactgtatggggcctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2552
 Db 185 GCACTGTATGGGGCTGGGCTTGTCTGTGGGGCTGGAATCTGGGTATGCTGTGTGACG 126
 QY 2593 aggtcagctgtgcaatcagtttaaacacccggaagaaacatttaacagcttatat 2652
 Db 125 AAGGTACGCTTACAAATCAAGTTAAACACCGAGAGAACATTTACATGCACTTATAT 66
 QY 2653 tctgtgtacacatctattctcaagcctaaaggtatgaagtgcctgtgtgtgtgtgt 2712
 Db 65 TTCTGTGTACATCTATTCTCAAAAGCTAAAGGTATGAAGTGCCTGCTTTGTATAG 6
 QY 2713 cca 2715
 Db 5 CCA 3

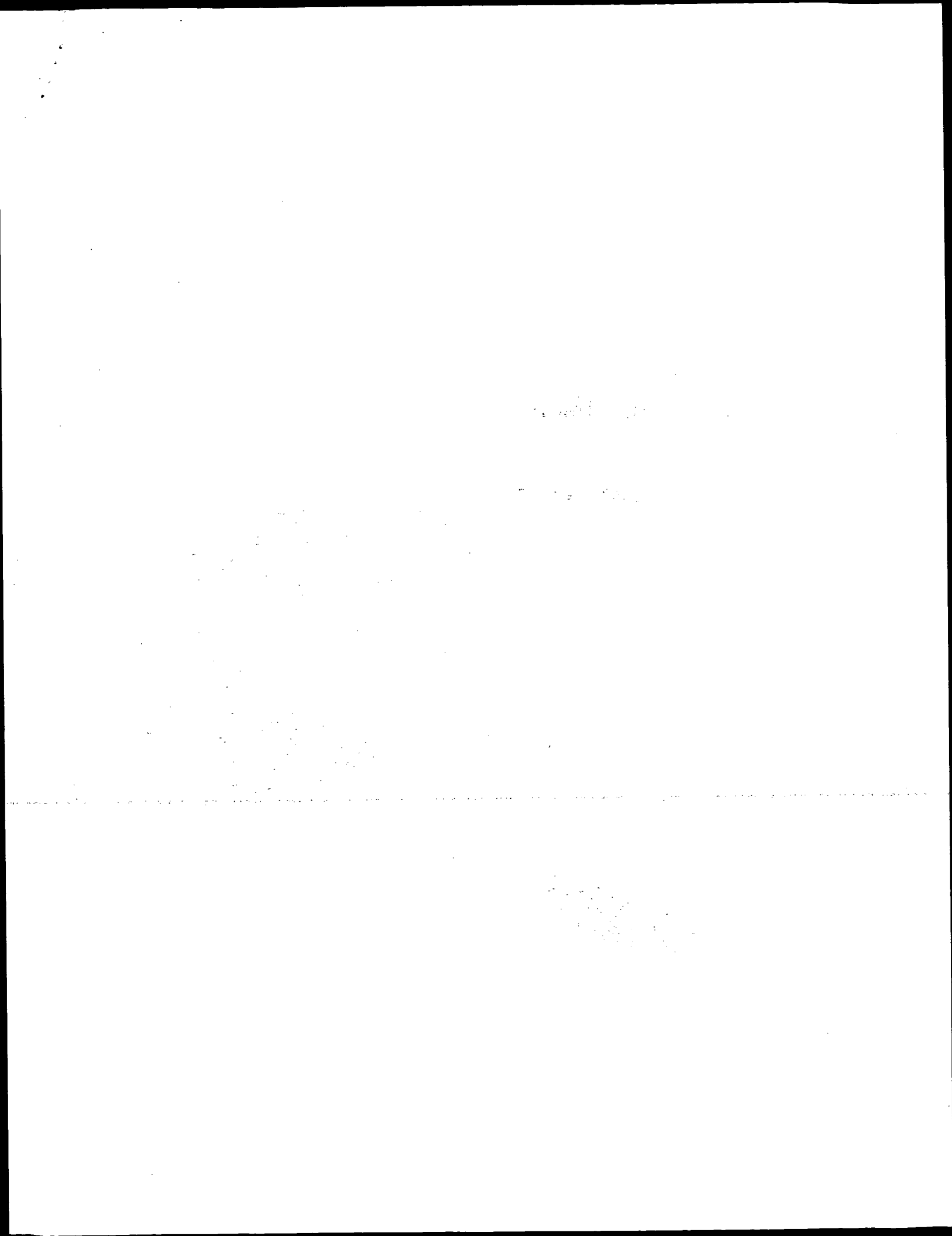
CONFIDENTIAL (SECRET)

Waterston, R. and Wilson, R.

Waterston, R. and Wilson, R.

QY	451	ttgttaagatgcgaaggccttttaagaaaggccatgaaaagcaagcccgctgaagttgc	510
Db	207	tgctaaagcctcgaaggcttcttaagccgaagcattgaaggccaagggccctgtttcaccctgc	266
Db	147	cctgggaatcttgagatgctgtggagaccgaaccaggcgctttccacttcaacgcattatacc	206
QY	391	cccccaatcttccttgatgtgaggaacaagaagccctctgcatcctacctcaattgtcatgaca	450

```
Search completed: February 18, 2001, 19:19:51
Job time: 17055 sec:
```



GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 18, 2001, 14:29:41 ; Search time 124.29 Seconds
(without alignments)
3766.761 Million cell updates/sec

Title: US-09-143-828-1

Perfect score: 2905
Sequence: 1 cctctgaaggtcttagatc.....aaaaaaaaaaaaaaaaaaaaa 2905

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 280836 seqs, 80580151 residues

Total number of hits satisfying chosen parameters: 561672

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: Issued_Patents_NA.*
2: /cgn2_6/ptodata/2/1na/5A.COMB.seq.*
3: /cgn2_6/ptodata/2/1na/5B.COMB.seq.*
4: /cgn2_6/ptodata/2/1na/6.COMB.seq.*
5: /cgn2_6/ptodata/2/1na/PCrus.COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	278.8	9.6	2043	1	US-07-737-736B-6
2	272.8	9.4	1399	1	US-07-737-736B-5
3	170.6	5.9	1450	1	US-08-458-489-1
4	170.6	5.9	1450	1	US-08-458-489-1
5	170.6	5.9	1450	1	US-07-843-350C-1
6	170.6	5.9	1450	4	US-08-342-411A-1
7	97.6	3.4	1898	1	US-08-648-619B-1
8	97.6	3.4	1979	2	US-08-648-619B-1
9	97.6	3.4	2030	1	US-08-330-518-1
10	97.6	3.4	2030	1	US-08-330-283-1
11	97.6	3.4	2030	2	US-08-646-248-1
12	97.6	3.4	2030	4	US-08-646-248-1
13	97.6	3.4	2030	4	US-08-646-248-1
14	97.2	3.3	1813	4	US-08-646-248-1
15	97.2	3.3	1813	4	US-08-646-248-1
16	97.2	3.3	1813	4	US-08-646-248-1
17	97.2	3.3	1813	4	US-08-646-248-1
18	97.2	3.3	1813	4	US-08-646-248-1
19	97.2	3.3	1813	4	US-08-646-248-1
20	97.2	3.3	1813	4	US-08-646-248-1
21	97.2	3.3	1813	4	US-08-646-248-1
22	97.2	3.3	1813	4	US-08-646-248-1
23	97.2	3.3	1813	4	US-08-646-248-1
24	97.2	3.3	1813	4	US-08-646-248-1
25	97.2	3.3	1813	4	US-08-646-248-1
26	97.2	3.3	1813	4	US-08-646-248-1
27	97.2	3.3	1813	4	US-08-646-248-1
28	97.2	3.3	1813	4	US-08-646-248-1

29	79.6	2.7	1659	1	US-08-333-358-7	Sequence 7, Appl
30	79.6	2.7	1659	1	US-08-463-694-7	Sequence 7, Appl
31	79.6	2.7	1659	1	US-08-694-501-7	Sequence 7, Appl
32	77.6	2.7	2970	4	US-08-694-501-7	Sequence 2, Appl
33	77.6	2.7	1688	2	US-08-649-619B-2	Sequence 1, Appl
34	74.8	2.6	1649	2	US-08-466-120-1	Sequence 1, Appl
35	74.8	2.6	1649	2	US-08-466-120-1	Sequence 1, Appl
36	74.4	2.6	1576	5	US-08-463-694-11	Sequence 1, Appl
37	74.4	2.6	2241	3	US-09-144-759-17	Sequence 19, Appl
38	74.4	2.6	2295	3	US-09-144-759-19	Sequence 21, Appl
39	74.4	2.6	2301	3	US-09-144-759-21	Sequence 11, Appl
40	73.2	2.5	2468	1	US-08-333-358-11	Sequence 11, Appl
41	73.2	2.5	2468	1	US-08-463-694-11	Sequence 11, Appl
42	73.2	2.5	2468	1	US-08-463-694-11	Sequence 11, Appl
43	72.6	2.5	1893	5	US-08-694-501-11	Sequence 11, Appl
44	68.6	2.4	1677	2	US-08-372-652-14	Sequence 14, Appl
45	68.6	2.4	1677	2	US-08-372-652-14	Sequence 14, Appl

ALIGNMENTS

RESULT 1
US-07-737-736B-6
Sequence 6, Application US/07737736B
Patent No. 5260199
GENERAL INFORMATION:
APPLICANT: Deluca, Hector F.
APPLICANT: Ross, Troy K.
APPLICANT: Prahl, Jean M.
TITLE OF INVENTION: Method Of Producing
TITLE OF INVENTION: 1,25-Dihydroxyvitamin D3 Receptor Protein
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Carl R. Schwartz, Esq., c/o Quarles & Brady
STREET: 411 East Wisconsin Avenue
CITY: Milwaukee
STATE: Wisconsin
COUNTRY: U.S.A.
ZIP: 53202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07737,736B
FILING DATE: 19910730
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Schwartz, Carl R.
REGISTRATION NUMBER: 29,437
TELEPHONE: 414-277-5715
TELEFAX: 414-277-5774
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 2043 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Rat
PUBLICATION INFORMATION:
AUTHORS: Bumester, James K.
AUTHORS: Wiese, Russell J.
AUTHORS: Maeda, No. 5260199uyuo
AUTHORS: Deluca, Hector F.
TITLE: Structure and regulation of the rat

Query Match	5.9%	Score 170.6	DB 1	Length 1450
Best Local Similarity	57.5%	Pred. No 3.3e-26		
Matches 327	Conservative 0	Mismatches 239	Indels 3	Gaps 1

QY	992	ccctgtctgccccacatgctgcacatgtaacctacatgcttcaagaagcatcatcagctt	1051
Db	739	ctctgtctacacacrrtgcgcacacatcaaaccttccatgctcgaagtcacatcaagttta	798
	1052	ccaaagtcacatcttctaacttcaggagcttgcacatcgaggaacagagatctccctcgtgaag	1111

Db	799	CTAAGGACCTGCGCCGCTTCCTCCGTCCTGACCATTTAAACACAGATCCCTTCTCAAG	858
QY	1112	gggcccgttcgagctgtgtcaactgaatcaacaagtgctcaacgcgggaacttgaa	1171
Db	859	GAGCAGCTGTGGAAATCTGTCACTCATGTACTCATATCCACTTCTGTCTCAAAACAAA	918
QY	1172	ccgtggaggtgtgcccgcctgtcctactgtctgtgaagaacactgcaggt--ggtccagc	1228
Db	919	ACTTCTCTCTGGGGCCCTTTCGCTACACATTTAAGATGTGAGACCCGTGTGGGCTTCAGG	978
QY	1229	aacttctactggaagcccaactgtcgaattccactacatgtctgaagaagctgcaactg	1288
Db	979	TAGAGTTTGTGAGACTTCTCTTTCACCTTCATGGAACACTAGCAAAATCGACGCTCCAG	1038
QY	1289	aggaagagatgtgtgtcatgtcagagcaatccctctcccccagaagcccggaagtgtgc	1348
Db	1039	AGCCTGATGATGTCCTCTTGAGCTGCTCCATGGCCCTGTCTCTCTGACCACTGAGATT	1098
QY	1349	tgcaagccgcggtgtgtgaccagctgcgaaggaacattcgccattactgtgaactctaca	1408
Db	1099	CCGACAGAGATGAGATTTGATGACGTGCAAGAGAGATGGACACTGACTGCAAACTGACA	1158
QY	1409	ttgaatgcacatgggcccagcgtgtctatagtgtcttctgtctgaagatcaatggtatgc	1468
Db	1159	TCAAGGGCCACACACGGAAGGCCCGGGGATGGTTCTGTATGGCAAGTGTCTAGGGCTGC	1218
QY	1469	tcaccagctccgagatcaatgtctcagcaacccagcggctgtctgcgcatccagagaca	1528
Db	1219	TGGCTGAGCTCCGAGCATTTAATGATGAGCCCTACGGGTATCAAAATCCAGCAATCCAGGGCC	1278
QY	1529	tacaccccttgcctagcggccctcatgtcag	1557
Db	1279	TGTCTGATGATGCTGCGCTGCTCCAGGAG	1307
RESULT	4		
US-08	458-686-1		
Sequence	1, Application	US/08458686	
Patent	No. 5710017		
GENERAL INFORMATION:			
APPLICANT:	David D. Moore et al.		
TITLE OF INVENTION:	CAR RECEPTORS AND RELATED		
TITLE OF INVENTION:	MOLECULES AND METHODS		
NUMBER OF SEQUENCES:	10		
CORRESPONDENCE ADDRESS:			
ADDRESSEE:	Fish & Richardson		
STREET:	225 Franklin Street		
CITY:	Boston		
STATE:	Massachusetts		
COUNTRY:	U.S.A.		
ZIP:	02110-2804		
COMPUTER READABLE FORM:			
MEDIUM TYPE:	3.5" Diskette, 1.44 MB		
COMPUTER:	IBM PS/2 Model 50Z or 55SX		
OPERATING SYSTEM:	IBM P.C. DOS (Version 3.30)		
SOFTWARE:	WordPerfect (Version 5.0)		
CURRENT APPLICATION DATA:			
APPLICATION NUMBER:	US/08/458,686		
FILING DATE:			
CLASSIFICATION:	514		
PRIOR APPLICATION DATA:			
APPLICATION NUMBER:	07/843,350		
FILING DATE:	February 26, 1992		
ATTORNEY/AGENT INFORMATION:			
NAME:	Paul T. Clark		
REGISTRATION NUMBER:	30,162		
REFERENCE/DOCKET NUMBER:	00786/126001		
TELECOMMUNICATION INFORMATION:			
TELEPHONE:	(617) 542-5070		
TELEFAX:	(617) 542-8906		
TELEX:	200154		
INFORMATION FOR SEQ ID NO: 1:			

```

;
; SEQUENCE CHARACTERISTICS
;
; LENGTH: 1450
;
; TYPE: nucleic acid
;
; STRANDEDNESS: single
;
; TOPOLOGY: linear
;
US-08-458-686-1

```

Query Match	5.98;	Score 170.6;	DB 1;	Length 1450;
Best Local Similarity	57.5%;	Pred. No. 3.3e-29;		
Matches 327; Conservative	0;	Mismatches 239;	Indels 3;	Gaps 1

QY	992	cctgtgtcccccacatgagctgtacatgtaactaactaagttaaagagatcaatgagctt	1051
Db	739	CTCTGGGTCACACACTTTCGGAGATCAATCAACATCTTATGTGACTGCAGATCATCAAGTTTA	798
QY	1052	ccaagatcatctctactactcaaggacttcccatcgagaccsaatctccctgtctgaag	1111
Db	799	CTAAGGACCTGCCCGCTTCCTCCGTTCCCTCCCATTTGAAGACCAAGATCTCCCTTCTCAAG	858
QY	1112	gggcccgtttcgaagctgtgtcaacttgaattcaacaasgtgttcaagcggagaacttggaa	1171
Db	859	GAGCAGCTGGGAAATCTGCACATGATGTACTCAAAACACATTTCTGTCTCCAAACACAAA	918
QY	1172	ccgtggaggtgtggccggctgtcctactgtcttggaagaacatgcaggt--ggtcttcagc	1228
Db	919	ACTTCTCTCTGGGGCTCTTCTGCTACACAAATGAAGAGAGACCCTGTGGGGTTTCCAG	978
QY	1229	aactctacatgagcccatctgtaaattccactatactgtgaagaagctgacgtcatg	1288
Db	979	TAGAGTTTGTGGAGTTGCTCTTTCATCTTCATCANGAAGACTAGAAATCGACAGCTCCAG	1038
QY	1289	aggaagagatctgtgtcatgtcagagcatctccctcttctcccaagaccggccaggtgtgc	1348
Db	1039	AGCCTGAGTATGTGCTCTTCTTGGCTGCATGAGCCCTGTCTCTCTACCGACCTGGAGTTA	1098
QY	1349	tgcagcacccggtgtgtgagaccagctgcagagaacattgcctacttgaagtccctaca	1408
Db	1099	CCCGAGAGAGATGAGATTTGATCACTGCTGMAAGAGATGCTACTGTCTGCAAAAGCTTCA	1158
QY	1409	ttgaatgcnaatccgccccacagctgcctcataggtcttctgttccctgaagaatcatgtctatgc	1468
Db	1159	TCAAGGGCCAGACAGCGAAGGCCCCCGGGAGATCGGTTTCTGTATCGAAGTTGTCAAGGCTTGC	1218
QY	1469	tcacagagcttcgcagcatcaatgtctcagacacaccagagcgtgtgtgcgtatccagagca	1528
Db	1219	TGGCTGAGACTCCGGAGCATTTAATGAAGGCTTACGGGTACCAAAATTCAGACACATCAAGGCC	1278
QY	1529	tacaaccttgtctcagagccctcatgcag 1557	
Db	1279	TGTCTGCATATATGCCCTGTCTCCAGAG 1307	

RESULT 5
 US-07-843-350C-1
 : Sequence 1 / Application US/07843350C
 : Patent No. 5756448
 : GENERAL INFORMATION:
 : APPLICANT: David D. Moore et al.
 : TITLE OF INVENTION: CAR RECEPTORS AND RELATED
 : TITLE OF INVENTION: MOLECULES AND METHODS
 : NUMBER OF SEQUENCES: 10
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Fish & Richardson
 : STREET: 225 Franklin Street
 : CITY: Boston
 : STATE: Massachusetts
 : COUNTRY: U.S.A.
 : ZIP: 02110-2804
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: 3.5" Diskette 1.44 Mb
 : COMPUTER: IBM PS/2 Model 50Z or 55SX
 : OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)


```

Db 417 CAGACTGGGTATCCAGATCCGAGAGAACAGACGCGCAAGGAGAAAGGCGCCAG 476
QY 374 atgagaagtcggaagtcgcccaatctgcgtgtatgtgaggaagcagcctgtatc 433
Db 477 CCCCAGAGATGCTGGGCGACAGACTTGGCTGTCTGTGGGACAAAGGCTCGGCTTCC 536
QY 434 actcaatgtcatatgacatgtgaagatgcaagagcttttcaagagggccatgaagca 493
Db 537 ACTACAACGTCTCAGCTCGAGAGCTGCAAGGGCTTCTCCGGCCAGTGTGTCCTGG 596
QY 494 agcccggtgtgagtgcccttcggaagggc---cctgagatcacccggaagacc 550
Db 597 GTGGGCGCAGGCGCTATGCTGCGGGGCTGGCGAAGCTGCCAGATGACGCTTTCATGC 656
QY 551 ggcagacgtgcagagcctgcgcctgcgaagtctggaagagcgagcatgaagagaga 610
Db 657 GCGGCAAGTCCAGCAGATGCGCGCTGCGCAAGTCAAGGAGGAGATGAGGAGCAGT 716
QY 611 tgatcatgtcgaagagcgccgtgagagagagcgcgcttgatcaagcggaagaaagt 670
Db 717 GCGTCTTCTCTGAAGAACAGATCCGAGAGAAAGATTGGAAGAACAGCAGCAGAGT 776
QY 671 aagcagagagactcagccactgtggaagtgcag 702
Db 777 CACAGTCACAGTCGACATCACTGTGGGCGC 808

```

RESULT 10
US-08-330-283-1
Sequence 1, Application US/08330283

Patent No. 5679518

GENERAL INFORMATION:

APPLICANT: Friedman, Eitan

APPLICANT: Holloway, M. Katharine

APPLICANT: Rodan, Gideon

APPLICANT: Rutledge, Su Jane

APPLICANT: Schmidt, Azriel

TITLE OF INVENTION: METHOD FOR FINDING RECEPTOR POTENTIATORS

NUMBER OF SEQUENCES: 5

CORRESPONDENCE ADDRESS:

ADDRESSEE: Merck & Co., Inc.

STREET: 126 East Lincoln Avenue

CITY: Rahway

STATE: New Jersey

COUNTRY: US

ZIP: 07065-0907

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/330,283

FILING DATE:

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Dolan, Catherine A.

REGISTRATION NUMBER: 36,502

REFERENCE/DOCKET NUMBER: 19327

TELECOMMUNICATION INFORMATION:

TELEPHONE: (908) 594-4283

TELEFAX: (908) 594-4720

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 2030 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

HYPOTHETICAL: NO

ANTI-SENSE: NO

US-08-330-283-1

Query Match 3.48; Score 97.6; DB 1; Length 2030;
Best Local Similarity 54.88; Pred. No. 7,1e-13;
Matches 215; Conservative 0; Mismatches 174; Indels 3; Gaps 1;

```

QY 314 ctgactgttactgtgagagacagatctgttccgtgaagccagtgtaacgcag 373
Db 417 CAGACTGGGTATCCAGATCCGAGAGAACAGACGCGCAAGGAGAAAGGCGCCAG 476
QY 374 atgagaagtcggaagtcgcccaatctgcgtgtatgtgaggaagcagcctgtatc 433
Db 477 CCCCAGAGATGCTGGGCGACAGACTTGGCTGTCTGTGGGACAAAGGCTCGGCTTCC 536
QY 434 actcaatgtcatatgacatgtgaagatgcaagagcttttcaagagggccatgaagca 493
Db 537 ACTACAACGTCTCAGCTCGAGAGCTGCAAGGGCTTCTCCGGCCAGTGTGTCCTGG 596
QY 494 agcccggtgtgagtgcccttcggaagggc---cctgagatcacccggaagacc 550
Db 597 GTGGGCGCAGGCGCTATGCTGCGGGGCTGGCGAAGCTGCCAGATGACGCTTTCATGC 656
QY 551 ggcagacgtgcagagcctgcgcctgcgaagtctggaagagcgagcatgaagagaga 610
Db 657 GCGGCAAGTCCAGCAGATGCGCGCTGCGCAAGTCAAGGAGGAGATGAGGAGCAGT 716
QY 611 tgatcatgtcgaagagcgccgtgagagagagcgcgcttgatcaagcggaagaaagt 670
Db 717 GCGTCTTCTCTGAAGAACAGATCCGAGAGAAAGATTGGAAGAACAGCAGCAGAGT 776
QY 671 aagcagagagactcagccactgtggaagtgcag 702
Db 777 CACAGTCACAGTCGACATCACTGTGGGCGC 808

```

RESULT 11
US-08-646-248-1

Sequence 1, Application US/08646248

Patent No. 5939322

GENERAL INFORMATION:

APPLICANT: Friedman, Eitan

APPLICANT: Holloway, M. Katharine

APPLICANT: Rodan, Gideon

APPLICANT: Rutledge, Su Jane

APPLICANT: Schmidt, Azriel

TITLE OF INVENTION: METHOD FOR FINDING RECEPTOR POTENTIATORS

NUMBER OF SEQUENCES: 5

CORRESPONDENCE ADDRESS:

ADDRESSEE: Merck & Co., Inc.

STREET: 126 East Lincoln Avenue

CITY: Rahway

STATE: New Jersey

COUNTRY: US

ZIP: 07065-0907

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/646,248

FILING DATE: 14-MAY-1996

CLASSIFICATION: 536

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/330,283

FILING DATE: 27-OCT-1994

ATTORNEY/AGENT INFORMATION:

NAME: Dolan, Catherine A.

REGISTRATION NUMBER: 36,502

REFERENCE/DOCKET NUMBER: 19327

TELECOMMUNICATION INFORMATION:

US-08-666-248-1

Query Match	3.48;	Score	97.6;	DB 2;	Length	2030;			
Best Local Similarity	54.80;	Pred. No.	7.1e-13;						
Matches	215;	Conservative	0;	Mismatches	174;	Indels	3;	Gaps	1

OY 334 ctgactctgttccacactcgttgaagacaaagagctctgttcttcgtgaagcccagtgtaacagcag 3737
 Db 417 CAGACTGGGTGATCTCCCAATCCCGAAGGAGAACCAAGAGCCGCAACCGAAGAAAGGCCAG 4767
 OY 374 atgaggaagtcggaaggtcccaaatctgcgttgatgtatgtggagacaagagccatgtgatac 4333
 Db 477 CCCCAGAAATGCTGGGGCAGCAGAGCTTTTCCGTGTGTGTGGGGACAAGAGCCCTCCGGCTTCC 5366
 OY 434 acttcaatgtcatgacatgtgaagaaatgcaaaagagccttttcaagaaagagccatgaaacgca 4933
 Db 537 ACTAACAGGTCTCAAGCTGCAGAAAGGCTGCAGAGGGCTCTTCCGGCGCAGTGTGGTCCGTG 5366
 OY 494 acgcccgcgtgaagtggtcccttcctcggaagagcg---cctcgagaatacaccggaagagccc 5530
 Db 597 GTGGGGCCAGGCGCCTATGCTCTGCGGGGTGGCGGAACCTGCACATGACAGCTTTCTATGC 6566
 OY 551 ggcgaacagtgcgaagcctgtccgcctctgcgaagatgcctcggaagagcgcgatgaagaagaga 6101
 Db 657 GGGCGAAGTGCACAGAGTGGCGGTGGCGCAAGTGCAGAGAGCGCAGATAGGAGCAGT 7166
 OY 611 tgatcatgtccgaacgagcgctggaagaaagagcgcgcccttgatcaagcggaaagaaagt 6701
 Db 717 GCGGTCTTTTGAAGAACAGATCCGGAAAGAAAGATTCCGAACACACACACAGAGAGT 7766
 OY 671 aacgagacgaagcctcagccaatcagtgagtgacg 702
 Db 777 CACAGTCACAGTCCGAGTCACTTGTGGGGCGG 808

RESULT 12
PCT-US95-13924-1
Sequence 1, Application PC/TUS9513924
GENERAL INFORMATION:
APPLICANT: Friedman, Eitan
APPLICANT: Holloway, M. Katharine
APPLICANT: Rodan, Gideon
APPLICANT: Rutledge, Su Jane
APPLICANT: Schmidt, Azriel
APPLICANT: Vogel, Robert
TITLE OF INVENTION: METHOD FOR FINDING RECEPTOR POTENTIALS
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Merck & Co., Inc.
STREET: 126 East Lincoln Avenue
CITY: Rahway
STATE: New Jersey
COUNTRY: US
ZIP: 07065-0907
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/13924

```

1      FILLING DATE:
2
3      CLASSIFICATION:
4
5      ATTORNEY/AGENT INFORMATION:
6
7      NAME: Oneag1@co, Carol S.
8      REGISTRATION NUMBER: 35,730
9      REFERENCE/DOCKET NUMBER: 15327 PCT
10
11     TELECOMMUNICATION INFORMATION:
12
13     TELEPHONE: (908) 594-3809
14     TELEFAX: (908) 594-4720
15
16     INFORMATION FOR SEQ ID NO.: 1:
17
18     SEQUENCE CHARACTERISTICS:
19
20     LENGTH: 2030. Base pairs
21
22     TYPE: nucleic acid
23
24     STRANDEDNESS: single
25
26     TOPOLOGY: linear
27
28     MOLECULE TYPE: DNA (genomic)
29
30     HYPOTHETICAL: NO
31
32     ANTI-SENSE: NO
33
34     PCT-US95-13924-1

```

Query Match	3.48	Score 97.6	DB 4	Length 2030
Best Local Similarity	54.88	Pred. No. 7.1e-13		
Matches 215; Conservative	0	Mismatches 174	Indels 3	Gaps 1

QY	314	ctactctttcaacacgttgaagacacaaagatcttcttccttgaaagcccgagtgcacagag	372
Db	417	CAGACGTGGGTCAATCCCAAGATCCCGAAGAGGAAACCAAGACGCCAAGCAAGAAAGGGCCAG	478
QY	374	atgaagaaatcggagagtcaccaaatctgcgctgataitgaggacaagaccatggtatc	433
Db	477	CCCCGAAGATGCTGGGCCACAGACGCTTTGGCGTGCTGTGTGGGGCAAGGCGTCCGGCTCC	538
QY	434	acttcaatgtctatgcacgtgaagaaagacgaagggcttttcaggaagggccatgaagaagca	495
Db	537	ACTACAAAGTGCTCAAGCTGCGCAAGGCGTCAAGGGCTTCTTCGGCGCGCAAGTGATGCTCGTG	598
QY	494	acgcgccggctcggagtggtgcccttccttcggaaagagcg-----ctcgcgagatctcaaccggaaagccc	550
Db	597	GTGGGGCCAGAGGGCGTAACTCTCTGCGGGGTGGCCGAAACCTTGCAAGATGAGACGCTTTCAGTC	656
QY	551	ggcgcaagatgycacagcctctgcgcgcctctgcgcgaatgtccttgaagagcgcataagaaagagaga	610
Db	657	GGCGCAAGTGCACAGACAGTGGCGGCTGGCGCAAGTGCACAGGAGGACAGGATGAGGGAGCGACT	716
QY	611	tgtatcatgtccgaacgaagcgcttgaagaaagagcgcggtccttgatcaagcggaagaagaatg	670
Db	717	GCGCTCTTTCTGTAAMACAGATCCGGAAAGAAAGATGTCGAAACAGCAGCAGCAGAGT	778
QY	671	aaagcagagagacatcagccaaatgtagtgtag	702
Db	777	CACAGTCACAGTCCAGTCCACTCTGGGGCCG	808

RESULT 13
 PCT-US95-13931-1
 : Sequence 1, Application PC/TUS9513931
 :
 : GENERAL INFORMATION:
 :
 : APPLICANT: Friedmann, Eitan
 : APPLICANT: Holloway, M. Katharine
 : APPLICANT: Rodan, Gideon
 : APPLICANT: Schmidt, Azriel
 : APPLICANT: Vogel, Robert
 :
 : TITLE OF INVENTION: USE OF RECEPTOR POTENTIATORS
 :
 : NUMBER OF SEQUENCES: 5
 :
 : CORRESPONDENCE ADDRESS:
 : ADDRESS: Merck & Co., Inc.
 : STREET: 126 East Lincoln Avenue
 : CITY: Rahway
 : STATE: New Jersey
 : COUNTRY: US
 :
 : ZIP: 07065-0907
 :
 : COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: PCT/US95/13931
 FILING DATE:
 CLASSIFICATION:
 ATTORNEY/AGENT INFORMATION:
 NAME: Quagliato, Carol S.
 REGISTRATION NUMBER: 35,330
 REFERENCE/DOCKET NUMBER: 19316 PCT
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (908) 594-4720
 TELEFAX: (908) 594-4720
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2030 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 HYPOHETICAL: NO
 ANTI-SENSE: NO
 PCT-US95-13931-1

Query Match 3.4%; Score 97.6; DB 4; Length 2030;
 Best Local Similarity 54.8%; Pred. No. 7.1e-13;
 Matches 215; Conservative 0; Mismatches 174; Indels 3; Gaps 1;

QY 314 ctgacttctgacactgtgagacacagagctgtctctgtaagcccaagtgtcaagcag 373
 DB 417 CAGACTGGGTATCCAGATCCAGAGAGAACAGAGCGACGACGAGAGAGGCGCCAG 476
 QY 374 atgaggaagtgcgggtcccaaatctgctgtatgtgggagaaagccactgtatc 433
 DB 477 CCCCCAATGCTGGGCGACGAGCTTTCGCTGTCTGTGGGAGAAAGGCTCCGCTTCC 536
 QY 434 actcaatgatacatgacatgtgagagatgcaagagcttttcaaggagggccatgaaagca 493
 DB 537 ACTCAAGTGTCTGACGTGCGAAGGCTGCAAGGGCTTTCGCGGCGAGTGTGCTCGTG 596
 QY 494 acgcccgtcgtagtgcccttcggaagggc---cctggaatatacccccgaagccc 550
 DB 597 GTGGGGCAGGCGCTATCTGCTGCGGGGTGGCGAAGCTGCAAGATGAGCCCTTTCATGC 656
 QY 551 ggcgcagctgacagcctgcgcctgcgcgaagtgcctggaagcgcgtgaaagagaga 610
 DB 657 GCGCGAAGTGCAGCAGATGCGCGCTGCGCAAGTGCAGAGGAGGAGGAGGAGGAGCAGT 716
 QY 611 tgatcatgtccgacgagcgttggagagagcgggacctgatacaagcggagaagaagt 670
 DB 717 GCGTCTTTTGAAGAACAGATCCGGAAGAGAAAGATTCGGAACAGCAGCAGCAGAGT 776
 QY 671 aacggacaggaactcagccactgtggagtgcag 702
 DB 777 CACAGTCAAGTCCGACGTCACCTGTGGGGCG 808

RESULT 14
 PCT-US94-12883-3
 Sequence 3, Application PC/TUS9412883
 GENERAL INFORMATION:

APPLICANT:
 TITLE OF INVENTION: UBIQUITOUS NUCLEAR RECEPTOR: COMPOSITIONS AND
 TITLE OF INVENTION: METHODS
 NUMBER OF SEQUENCES: 38
 CORRESPONDENCE ADDRESS:
 ADDRESS: Arnold, White & Durkee
 STREET: P. O. Box 4433
 CITY: Houston
 STATE: Texas

COUNTRY: United States of America
 ZIP: 77210
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS/ASCII
 SOFTWARE: Patentin Release #1.0, Version
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: PCT/US94/12883
 FILING DATE: Concurrently Herewith
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/152,003
 FILING DATE: 10-NOV-1993
 CLASSIFICATION:
 ATTORNEY/AGENT INFORMATION:
 NAME: BARBARA S. KITCHELL
 REGISTRATION NUMBER: 33,928
 REFERENCE/DOCKET NUMBER: ARCD154P--
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (512) 418-3000
 TELEFAX: (713) 789-2679
 TELEX: 79-0924
 INFORMATION FOR SEQ ID NO: 3:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1813 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 PCT-US94-12883-3

Query Match 3.3%; Score 97.2; DB 4; Length 1813;
 Best Local Similarity 58.8%; Pred. No. 8.3e-13;
 Matches 187; Conservative 0; Mismatches 128; Indels 3; Gaps 1;

QY 388 gtctcccaaatctgcgtatgtgagacagcagcactggtatcattcaatgtcat 447
 DB 229 GCGCAGAGCTTGGCGTCTGTGTGGGACAGGCTTCGCGCTTCACTACAGCTGTC 288
 QY 448 acatgtgaagatgcaagagcttttcaaggagggccatgaaacgaaagcccgctgagg 507
 DB 289 AGCTGGAGAGGCTGCAAGGCTTTCGCGCGCAGATGTGCTCGGAGGCGCAGGCGC 348
 QY 508 tgccttcggaaggg---cctggaatatacccccgaagcccgcagtgagcag 564
 DB 349 TATGCTGCGGGGTGCGCAAGCTGCGAAGTGAAGCTTTCATGCGGCGCAAGTGCAG 408
 QY 565 gctgcgcctgagcagagtcctggaagcgcgcagcagcagcagcagcagcagcagcag 624
 DB 409 CAGTGGCGGCTGCGCAAGTGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 468
 QY 625 gaagcgttggagagagcggccttgatcaagcggagaagaagtgaacggagagagact 684
 DB 469 GAGCAGATCCGGAAGAAAGATTCGGAACAGCAGCAGCAGCAGCAGCAGCAGCAGT 528
 QY 685 cagccactggagtgagcag 702
 DB 529 CAGTCAACTGTGGGGCG 546

RESULT 15
 US-08-095-728B-3
 Sequence 3, Application US/08095728B
 GENERAL INFORMATION:

APPLICANT: DMITROVSKY, ETHAN
 APPLICANT: MARRELL JR, RAYMOND P
 APPLICANT: MILLER JR, WILSON H
 APPLICANT: FRANKEL, STANLEY
 TITLE OF INVENTION: METHODS FOR THE DETECTION AND

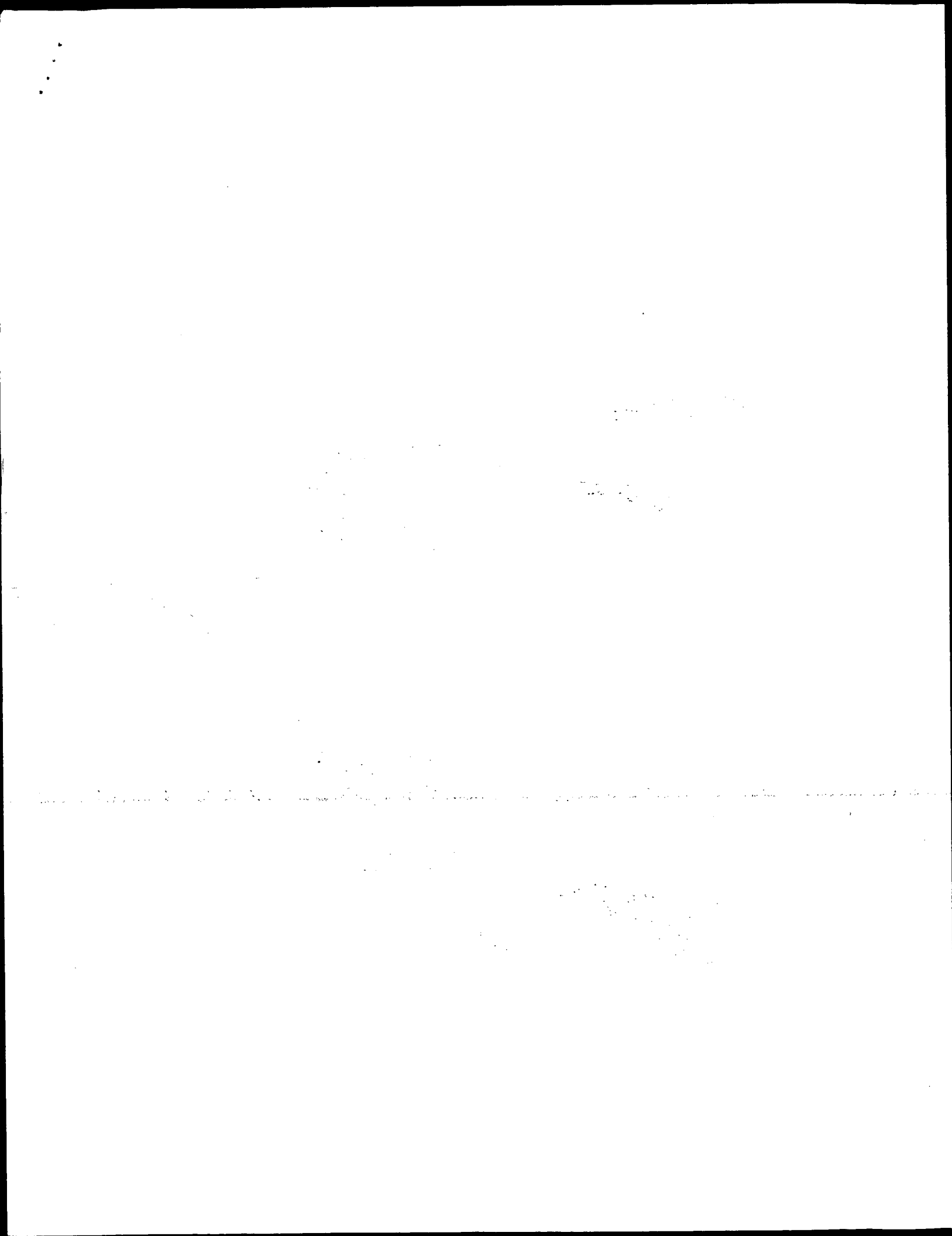
TITLE OF INVENTION: TREATMENT OF ACUTE PROMYELOCYTIC LEUKEMIA (APL)
 NUMBER OF SEQUENCES: 11
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: COOPER & DUNHAM LLP
 STREET: 1185 AVENUE OF THE AMERICAS
 CITY: NEW YORK
 STATE: NEW YORK
 COUNTRY: U.S.A.
 ZIP: 10036
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.24
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/095,728B
 FILING DATE: 21-JUL-1993
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/673,838
 FILING DATE: 22-MAR-1991
 ATTORNEY/AGENT INFORMATION:
 NAME: WHITE, JOHN P
 REGISTRATION NUMBER: 28,678
 REFERENCE/DOCKET NUMBER: 38694-A
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (212) 278-0400
 TELEFAX: (212) 391-0525
 INFORMATION FOR SEQ ID NO: 3:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2928 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: CDNA
 IMMEDIATE SOURCE:
 CLONE: hRAR ALPHA
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 103..1488
 OTHER INFORMATION:
 US-08-095-728B-3

Query Match 3.1%; Score 91.4; DB 2; Length 2928;
 Best Local Similarity 50.8%; Pred. No. 2e-11;
 Matches 302; Conservative 0; Mismatches 281; Indels 12; Gaps 3;

QY 993 cctgctgcccccatgctgctacatgctacatcatcattcaaggcatcatcagcttgc 1052
 DB 771 cctctggagacattcagtgacatctccacacagatcattttagactgtggagttccg 830
 QY 1053 caaagtcattctcctacttcagggagcttgcacatcgagagacagatctccctgctgaagg 1112
 DB 831 caagcagctgcccggcttgcacacccatccacacatccgacacagatcaccctctcaggcc 890
 QY 1113 ggcgccttcgagctgctgacatgagatcaaacacagtggtcaacgctgagactggaac 1172
 DB 891 tgcctgctgagacatctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 950
 QY 1173 ctggagtggt---ggcgcgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 1229
 DB 951 catgaccttctgagacggctgacccctgacacggacccagatccacacacgctgctgctgctg 1010
 QY 1230 acttctactgagagccatgctgaatctcactacatcagctgagaaagctgacagctgcatga 1289
 DB 1011 cccctctacacgacatgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 1070
 QY 1290 gggaggtatgtgtatgctgagagcattctcctctctctcctcctcctcctcctcctcctcct 1349
 DB 1071 tggggagaggggctgctgacgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 1130
 QY 1350 ggcagcagcgctgtgtgacacagctgacagagacattcgcttacttgaagtcctact 1409

DB 1131 gcagccggagacccggctgagacatctcctgagagccgctgctgagagcgcttgaagcttacct 1190
 QY 1410 tgaatgcaatcgcccccagcctgctcattagttcttcttctgtaagatcatggtatgct 1469
 DB 1191 gcggagacggagcccgccgccccacac-----tggtccccaagatgcttaagatg 1244
 QY 1470 caccgagctcgagacatcattatgctcagacacacccagcgctgct---ggcattccaga 1526
 DB 1245 tactgacctgccaagatcagcggccgaaggggctgagcggtgatcacctgaagatgga 1304
 QY 1527 catcaccccttctgtaagccctcattcagagagtgcttgcgcatcacaggtagc 1581
 DB 1305 gatccgggcttcattgcccgcctctcattcagagaaattgttgagaaactcagagggc 1359

Search completed: February 18, 2001, 16:49:16
 Job time: 8375 sec



RESULT	1
ID	X56242 standard; cDNA; 2910 BP.
AC	X56242;
DT	16-JUL-1999 (first entry)
DE	Human vitamin D receptor related gamma protein encoding cDNA.
XX	
XX	Human; vitamin D receptor related protein; VDR; obesity; diabetes;
XX	anorexia; rheumatoid arthritis; lipoprotein defect; hyperlipidaemia
XX	hypercholesterolaemia; hyperlipoproteinaemia; osteoporosis; tumour;
XX	hyperproliferative skin disorder; hyperthyroidism; ss.
OS	Homo sapiens.
XX	
PN	W09919354-A1.
XX	
PD	22-APR-1999.
XX	
PF	31-AUG-1998; 98WO-SE01548.
XX	
PR	31-MAR-1998; 98SE-0001148.
PR	14-OCT-1997; 97SE-0003745.
XX	
PA	(PHAA) PHARMACIA & UPJOHN AB.
XX	
PI	Berkenstam A, Dahlberg M;
XX	
WP	1999-302508/25.
DR	P-PsDB; Y09515.
XX	
PT	New vitamin D receptor related (VDR) polypeptides, useful for
XX	treating obesity, diabetes, anorexia and rheumatoid arthritis

XX Claim 2; Page 17-18; 35pp; English.

XX The present sequence encodes a human vitamin D receptor related (VDR) polypeptide. Human VDR polypeptides and substances which affect VDR signal transduction, can be used for treating metabolic, proliferative or inflammatory conditions. They can be used in the manufacture of a medicament for treating the following conditions: obesity, diabetes, anorexia, lipoprotein defects, hyperlipidemia, hypercholesterolemia or hyperlipoproteinemia and osteoporosis, rheumatoid arthritis, benign and malignant tumours, hyperproliferative skin disorders or hyperthyroidism. Nucleic acid vectors encoding for expression of a VDR polypeptide can be used for treating metabolic, proliferative or inflammatory conditions, by introducing them into a mammal. The introduced nucleic acid is then capable of transforming a cell in vivo and then polypeptide is expressed. A substance affecting VDR signal transduction, such as an agonist or antagonist can be used for the manufacture of a medicament for treating metabolic, proliferative or inflammatory condition.

CC N.B. The specification specifically claims the VDR nucleic acid and polypeptide sequences given in figures 1, 4, 7 and 8, but no figures are given in the specification.

XX Sequence 2910 BP; 765 A; 728 C; 780 G; 636 T; 1 other;

Query Match 99.6%; Score 2894; DB 20; Length 2910;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2905; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 cctctgaaggttcagatcagatagatgaattcgtcgagcggagagagagacatgcctt 60
Db 1 cctctgaaggttcagatcagatagatgaattcgtcgagcggagagagagacatgcctt 60
QY 61 tacttcagtggaatctcgcctcagcctcagcagccaagtgttcaacagtgaaaaagaa 120
Db 61 tacttcagtggaatctcgcctcagcctcagcagccaagtgttcaacagtgaaaaagaa 120
QY 121 gagataagcctaatactcctgtctctgaacagagcagcgcctctgttaagcctac 180
Db 121 gagataagcctaatactcctgtctctgaacagagcagcgcctctgttaagcctac 180
QY 181 tgatcagatccttgcacccgagatgttcaaaagtgaaccacagagagagagagtcgagcaag 240
Db 181 tgatcagatccttgcacccgagatgttcaaaagtgaaccacagagagagagagtcgagcaag 240
QY 241 aacttacccaacagcagtcgaagagcccaagaaacactgtgaggtgaaacccaagaa 300
Db 241 aacttacccaacagcagtcgaagagcccaagaaacactgtgaggtgaaacccaagaa 300
QY 301 agctggaacacatgtgactgtgtacactgtgtgagacacagaggtctgttccctggaagccc 360
Db 301 agctggaacacatgtgactgtgtacactgtgtgagacacagaggtctgttccctggaagccc 360
QY 361 agtgcacacagcagatgagagagtcgagagtcgcccaaaatctgcgtgtatgttgagagaaag 420
Db 361 agtgcacacagcagatgagagagtcgagagtcgcccaaaatctgcgtgtatgttgagagaaag 420
QY 421 ggcacatgctatcacttaatgtcatgacatgtgaaagagatgcaagagccttttcaagagag 480
Db 421 ggcacatgctatcacttaatgtcatgacatgtgaaagagatgcaagagccttttcaagagag 480
QY 481 ggcattgaaagcgaacgcccggccttgaggtgccccttcggaagagcgcctctggaatatacc 540
Db 481 ggcattgaaagcgaacgcccggccttgaggtgccccttcggaagagcgcctctggaatatacc 540
QY 541 cggagaaagcccgagacagtcgacagcctgcgcctgcgcaaggtgctcggagagagcgatg 600
Db 541 cggagaaagcccgagacagtcgacagcctgcgcctgcgcaaggtgctcggagagagcgatg 600
QY 601 aagaaagagatgatactgacgaagagcgctgagagagagcgagccttgataaagcgg 660
Db 601 aagaaagagatgatactgacgaagagcgctgagagagagcgagccttgataaagcgg 660

QY 661 aagaaagatgacagcagacagagcactcagccactggagagtgcaggggtctacagagagcaag 720
Db 661 aagaaagatgacagcagacagagcactcagccactggagagtgcaggggtctacagagagcaag 720
QY 721 cggatgatgatacaggaagcctgatgagacgtcagatgaataaccccttgacatactctcc 780
Db 721 cggatgatgatacaggaagcctgatgagacgtcagatgaataaccccttgacatactctcc 780
QY 781 cattcaagaattcccgctcgcagaggggtgtcttagcaatgctcgaggttcagagttc 840
Db 781 cattcaagaattcccgctcgcagaggggtgtcttagcaatgctcgaggttcagagttc 840
QY 841 ctgcagagcccatcagaggaagaaagcctcgaagtgaagccaggtcccggaagacatgtgc 900
Db 841 ctgcagagcccatcagaggaagaaagcctcgaagtgaagccaggtcccggaagacatgtgc 900
QY 901 tctttgaaggtctctctcagcgtcgggggagagatgagtcagtgctcgtgaatacaaaccc 960
Db 901 tctttgaaggtctctctcagcgtcgggggagagatgagtcagtgctcgtgaatacaaaccc 960
QY 961 ccagccgaagatgctgcgggaaagaaatctctccctgcgtccccaatgctgacatgtca 1020
Db 961 ccagccgaagatgctgcgggaaagaaatctctccctgcgtccccaatgctgacatgtca 1020
QY 1021 acctacatgtcgaagagcactacagccttgcgaagtcacatcctacttcaggagacttg 1080
Db 1021 acctacatgtcgaagagcactacagccttgcgaagtcacatcctacttcaggagacttg 1080
QY 1081 cccatcgaagagcagacatcctcctgtctgaaggggcgcttcagagctgtgtcaactgaga 1140
Db 1081 cccatcgaagagcagacatcctcctgtctgaaggggcgcttcagagctgtgtcaactgaga 1140
QY 1141 ttcacaacacgtttcaacggcgaagacatggaacacctgtgaggtgtgctgcggcttccctacgc 1200
Db 1141 ttcacaacacgtttcaacggcgaagacatggaacacctgtgaggtgtgctgcggcttccctacgc 1200
QY 1201 ttggaagacacatgcaggtgtgtcttcagcaacttctacagggcccatctgaataatccac 1260
Db 1201 ttggaagacacatgcaggtgtgtcttcagcaacttctacagggcccatctgaataatccac 1260
QY 1261 taacatgtgaaagacatgacagctgacatgagagagatgtgtgatagagagcaatctcc 1320
Db 1261 taacatgtgaaagacatgacagctgacatgagagagatgtgtgatagagagcaatctcc 1320
QY 1321 cctcttcctccagacagcccaaggtgtgtctgagacacgcggtgtgagacacatgcaagag 1380
Db 1321 cctcttcctccagacagcccaaggtgtgtctgagacacgcggtgtgagacacatgcaagag 1380
QY 1381 caattccgacatcactgagatcctacatctgaatgaaatcgccccagcctgctcatag 1440
Db 1381 caattccgacatcactgagatcctacatctgaatgaaatcgccccagcctgctcatag 1440
QY 1441 ttctgttcc - tgaagatcaatgctatgtctcaacagagctcccgacagatcaatgtctagca 1499
Db 1441 ttctgttccgtgaaagatcaatgctatgtctcaacagagctcccgacagatcaatgtctagca 1500
QY 1500 caccacagcgcgtgcgtgcgtatcagagacatacaaccccttgcgaagccctcatgcaaga 1559
Db 1501 caccacagcgcgtgcgtgcgtatcagagacatacaaccccttgcgaagccctcatgcaaga 1560
QY 1560 gtgtgtcagcatcaacagatgagctgagcgcgtgccccttggtgtgacaacccctcgagagcagc 1619
Db 1561 gtgtgtcagcatcaacagatgagctgagcgcgtgccccttggtgtgacaacccctcgagagcagc 1620
QY 1620 cagaccagaagccctctgagccgcgcacatcccgagcgaagacagatggaactgccaagag 1679
Db 1621 cagaccagaagccctctgagccgcgcacatcccgagcgaagacagatggaactgccaagag 1680
QY 1680 ccgacaatgcccctgtgctgtctcccttagagaaattcgtctatgagagctgtgtaaga 1739
Db 1681 ccgacaatgcccctgtgctgtctcccttagagaaattcgtctatgagagctgtgtaaga 1740
QY 1740 ttctctcaggaagacatggtgtcccccaaccccgatctcagtgctgtagaggtgaagca 1799

Db	1741	ttctccagaggaagacatggtctgccccccaccccccaggttcagctctgtaggggtgagcgca	1800
Qy	1800	caagacttcacgttggagagatgcaactgcaactgtatggttcagagacatcatagagagcaaggtc	1859
Db	1801	caagactctcagctggagagatgagcaactgcaactgtatggttcagagacatcatagagagcaaggtc	1860
Qy	1860	gacctcttcctcttttaaaagggccgctgtgtcttgggggaaatccctctcagttcccatatagtc	1919
Db	1861	gacctcttcctctctttaaagggccctgtgtctctgggggaaatccctctcagttcccatatagtc	1920
Qy	1920	gtccaaagtgtgtgaaagggagaccaaagctggagaccaaagatagcgcatcttgggtgtctatggccaat	1979
Db	1921	gtccaaagtgtgtgaaagggagaccaaagctggagaccaaagatagcgcatcttgggtgtctatggccaat	1980
Qy	1980	acccaagcttctgtcgtctctccgtgaagctcttccattgtgtaacctctaaatagtcctgtctccca	2039
Db	1981	acccaagcttctgtcgtctctccgtgaagctcttccattgtgtaacctctaaatagtcctgtctccca	2040
Qy	2040	cttcccaactcgtcttccctctctctctccggaactgtcttgttgggtctcaagccctgtactcatc	2099
Db	2041	cttcccaactcgtcttccctctctctctccggaactgtcttgttgggtctcaagccctgtactcatc	2100
Qy	2100	ggcaggggtgacgtgagatctctgtggagcttccctccaaagagagatgtgaaagccaaagagcctgtca	2159
Db	2101	ggcaggggtgacgtgagatctctgtggagcttccctccaaagagagatgtgaaagccaaagagcctgtca	2160
Qy	2160	ccaaagtctcagaaagctctgagatagacctcatctccggggccaacatcatctctgtctctgtccatc	2219
Db	2161	ccaaagtctcagaaagctctgagatagacctcatctccggggccaacatcatctctgtctctgtccatc	2220
Qy	2220	catttgaacacatctatagagcactgtatagataggtagcctgtcgtctgtgggttatacagcatgtg	2279
Db	2221	catttgaacacatctatagagcactgtatagataggtagcctgtcgtctgtgggttatacagcatgtg	2280
Qy	2280	actcagagatagagatcccgagactcagagagtttaagtttaaaaaaaacaaacagaaacaaacaa	2339
Db	2281	actcagagatagagatcccgagactcagagagtttaagtttaaaaaaaacaaacagaaacaaacaa	2340
Qy	2340	acaaattctgtgatacaaaagggagaaataatgtatgtacaaagacagacaaaggaatttccctgt	2399
Db	2341	acaaattctgtgatacaaaagggagaaataatgtatgtacaaagacagacaaaggaatttccctgt	2400
Qy	2400	tgtgtgatactgagagcgtctgtatgtgcagagcactgtgggtacccaagtgtgaaggtctcccgagagacat	2459
Db	2401	tgtgtgatactgagagcgtctgtatgtgcagagcactgtgggtacccaagtgtgaaggtctcccgagagacat	2460
Qy	2460	gagagctgtgaaaggaagggagacaaactgcagcgtgtgaggtgtgtgtgtgtgtgtgtgtgtgtgt	2519
Db	2461	gagagctgtgaaaggaagggagacaaactgcagcgtgtgaggtgtgtgtgtgtgtgtgtgtgtgtgt	2520
Qy	2520	ggtaaggtctgtgttgcacactgtgacgtgggggcctgtgtgtctctccgggggcctggaaatgtcgtgt	2579
Db	2521	ggtaaggtctgtgttgcacactgtgacgtgggggcctgtgtgtctctccgggggcctggaaatgtcgtgt	2580
Qy	2580	atgctctctgtgacaaggtcactcgtgacacatcagttaaaacacacggagagaaacacatcttac	2639
Db	2581	atgctctctgtgacaaggtcactcgtgacacatcagttaaaacacacggagagaaacacatcttac	2640
Qy	2640	atgcaaccttaataattctcgtgtataacatctatcttccaaagcttaaaaggtcatgtgaaggtgct	2699
Db	2641	atgcaaccttaataattctcgtgtataacatctatcttccaaagcttaaaaggtcatgtgaaggtgct	2700
Qy	2700	ggcctgttttaaaagcactctgtgtgaaataaaatttttttggatcttccaaattccaaatatatctcta	2759
Db	2701	ggcctgttttaaaagcactctgtgtgaaataaaatttttttggatcttccaaattccaaatatatctcta	2760
Qy	2760	tataaaggaattcccaacacttaagaaactagtttcttgggaaatgtagccctgggttttaagtcca	2819
Db	2761	tataaaggaattcccaacacttaagaaactagtttcttgggaaatgtagccctgggttttaagtcca	2820
Qy	2820	aatacaaggcaaaaggaatttaataatgtacttcttggcttaaaaaaaadtaaaaaaaadtaaaaaaa	2879

```

Db      2821   atcgaagccaaaggaaatlaataactgtctttgccaacaaaaaaaaaaaaaaa 2880
Oy      2880   aaaaaaaaaaaaaaaaaaaaaa 2905
                               |||||
Db      2881   aaaaaaaaaaaaaaaaaaaaaa 2906

RESULT    2
X78807
ID        X78807 standard; cDNA, 3093 BP.
XX
AC        X78807;
XX
DT        06-SEP-1999 (first entry)
XX
DE        Human NNR7 cDNA.
KW        NR7; NR7-1; nuclear trans-acting receptor protein; human; regulator;
KW        identification; downstream target gene; cell proliferation;
KW        cell development; ss.
XX
OS        Homo sapiens.
XX
FH        Key          Location/Qualifiers
FT        CDS           .275..1676
                  /tag=a
                  /product= "NNR7"
                  /note= "Partial sequence, no start codon given"
FT
FT
FT
PN        MO9931129-A1.
PD
PF        24-JUN-1999.
PE
PR        11-DEC-1998;     98WO-US26364;
PR        14-OCT-1998;     98US-0104251.
PR        12-DEC-1997;     97US-0069401.
PA        (MERI ) MERCK & CO INC.
PI
PI        Chen F;
DR        WPI: 1999-405024/34.
DX        P-FSDB: Y25410.
PT
PX        DNA encoding human nuclear receptors NR7 and NR7-1
PS
PS        Claim 9; Fig 1A-C; 80pp; English.
CC
CC        This invention describes the isolation of the novel human nuclear
CC        receptors NR7 or NR7-1. The NR7 and NR7-1 proteins are useful in
CC        the identification of downstream target genes and ligands regulating its
CC        activity. The nuclear receptor is involved in the regulation of in vivo
CC        cell proliferation and/or cell development. The NR7 and NR7-1
CC        polynucleotides, expression vectors and host cells are useful for the
CC        recombinant production of the protein.
CQ
SQ        Sequence 3093 BP; 774 A; 786 C; 843 G; 690 T; 0 other;

Query Match              96.5%; Score 2804.6; DB 20; Length 3093;
Best Local Similarity    99.5%; Pred. No. 0;
Matches 2824; Conservative 0; Mismatches 14; Indels 1; Gaps 1

Oy      22   ataaggaattcgtaggcggagaagggaagcaactgaccttacttcagtggaaatctcggc 81
            | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      114   ataacctaatgaactagacggaaggaaggaagcacctgcccttacttcagtggaaatctcggc 173
            | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Oy      82   cttaacgcttgcaagcaagtggttcacagtgaaaaaagcaagagaataagctaactctcy 141
            | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      174   ctcaagcttgcaagcaagtggttcacagtgaganaaaagcaagagaataagctaactctcy 233
            | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Oy      142   tccctaacacaagcgacgcgtcctctggttaaagctaactccttgatcatccttgcaaccgga 201

```

|||||
Db 234 tccgtaacagcagcgctcctcttgtaagactactcctgatcgatcctcttgacccgga 293
QY 202 ttgttccaagtggagcccgagggagaaatcgagacaaagaacttaaccaacgaagctcca 261
Db 294 ttgttccaagtggagcccgagggagaaatcgagacaaagaacttaaccaacgaagctcca 353
QY 262 agagcccgagagacaaacttgagagtgagacccaagaagctggaacatgagctt 321
Db 354 agagcccgagagacaaacttgagagtgagacccaagaagctggaacatgagctt 413
QY 322 gtacacgtgagagacagagatctgttctcgtgaaagccagtgtaacagcagatgagaa 381
Db 414 gtacacgtgagagacagagatctgttctcgtgaaagccagtgtaacagcagatgagaa 473
QY 382 gtccgagagtgcccaaaatctccgctgtatgtggagacaaagccactgtctacatcaat 441
Db 474 gtccgagagtgcccaaaatctccgctgtatgtggagacaaagccactgtctacatcaat 533
QY 442 gtcatgacatgtgaaagtgacaaagctttctcagagagccatgaaacgcaacgcccgg 501
Db 534 gtcatgacatgtgaaagtgacaaagctttctcagagagccatgaaacgcaacgcccgg 593
QY 502 ctgaggtgcccccttcgcggaagggcgctctgcagagatcaaccggaagcccgagcaagtg 561
Db 594 ctgaggtgcccccttcgcggaagggcgctctgcagagatcaaccggaagcccgagcaagtg 653
QY 552 cagagctgcgcgcctgcggaagtgctcgtgagagggcgatgaaagaaagtgatcatctcc 621
Db 654 cagagctgcgcgcctgcggaagtgctcgtgagagggcgatgaaagaaagtgatcatctcc 713
QY 622 gacagagcgctgagagaaagcgcgctctgatacaagcggaagaaagtgaacgagacaggg 681
Db 714 gacagagcgctgagagaaagcgcgctctgatacaagcggaagaaagtgaacgagacaggg 773
QY 682 actcagccactgagagtgagagggcgctgacagagagagcagcgatgatcaagagagctg 741
Db 774 actcagccactgagagtgagagggcgctgacagagagagcagcgatgatcaagagagctg 833
QY 742 atgagacgctcagatgaaagaactttgacactactctctccacttcaagaatctccgctg 801
Db 834 atgagacgctcagatgaaagaactttgacactactctctccacttcaagaatctccgctg 893
QY 802 ccaggggtgcttaagcagtgctgagtgacagagtgctcgcagggcccaacgcagagaa 861
Db 894 ccaggggtgcttaagcagtgctgagtgacagagtgctcgcagggcccaacgcagagaa 953
QY 862 gaagctgcagagtgagagcgagctccggaagagatctgtctcttgaaagtgctctcag 921
Db 954 gaagctgcagagtgagagcgagctccggaagagatctgtctcttgaaagtgctctcag 1013
QY 922 ctgggggggggagtgagagtgctgagacatacaaaaccccgagcgatgaggggaa 981
Db 1014 ctgggggggggagtgagagtgctgagacatacaaaaccccgagcgatgaggggaa 1073
QY 982 gagatctctccctgcctgccccacatgagctgacatgtcaactcaatgcttaagagctc 1041
Db 1074 gagatctctccctgcctgccccacatgagctgacatgtcaactcaatgcttaagagctc 1133
QY 1042 atcagcttgcagaaagtcatctcctactcagggagctgcccacgcagagacatctcc 1101
Db 1134 atcagcttgcagaaagtcatctcctactcagggagctgcccacgcagagacatctcc 1193
QY 1102 ctgctgtaagggggcgcttctgagctgtgttaactgagattcaacaacaggttcaagcg 1161
Db 1194 ctgctgtaagggggcgcttctgagctgtgttaactgagattcaacaacaggttcaagcg 1253
QY 1162 gagactgaaacctggagtgctggcgagctgctcactgagcttggaagacactgagcgagc 1221
Db 1254 gagactgaaacctggagtgctggcgagctgctcactgagcttggaagacactgagcgagc 1313
QY 1222 ttccagaaacttactgagagccatgctgaattcaactacatgctgaaagagctgag 1281
|||||
Db 1314 ttccagaaacttactgagagccatgctgaattcaactacatgctgaaagagctgag 1373
QY 1282 ctgcataagagagatgtgctgagcagggccatctccctcttcccgagacggcca 1341
Db 1374 ctgcataagagagatgtgctgagcagggccatctccctcttcccgagacggcca 1433
QY 1342 ggtgtgctgagacacggcgctggagacagctgagagagcaaatctgcatactctgag 1401
Db 1434 ggtgtgctgagacacggcgctggagacagctgagagagcaaatctgcatactctgag 1493
QY 1402 tccatcatgtgatcaatcgccccaagcctgtctataggtctctgttccctgaagatcatg 1461
Db 1494 tccatcatgtgatcaatcgccccaagcctgtctataggtctctgttccctgaagatcatg 1553
QY 1462 gctatgtcacccagagctccgagacataatgtctcagacacacccagcgctgagcatc 1521
Db 1554 gctatgtcacccagagctccgagacataatgtctcagacacacccagcgctgagcatc 1613
QY 1522 caggaatacaccccttctgacagccctcactgacagaggttgctgacatacaagtagc 1581
Db 1614 caggaatacaccccttctgacagccctcactgacagaggttgctgacatacaagtagc 1673
QY 1582 tgaagcgctgcccccttggtgagacactccgagagagcagccacagacccctctgagcc 1641
Db 1674 tgaagcgctgcccccttggtgagacactccgagagagcagccacagacccctctgagcc 1733
QY 1642 gccactcccggtccaaagagatgagacatgacagagccagcaatgacgctgagctg 1701
Db 1734 gccactcccggtccaaagagatgagacatgacagagccagcaatgacgctgagctg 1793
QY 1702 tctccctagggtaattctctgataagacagctggtgacatctcctcagagagagacatgag 1761
Db 1794 tctccctagggtaattctctgataagacagctggtgacatctcctcagagagagacatgag 1853
QY 1762 ccccccaccccaagctcagctgtgagagtgagagccacagactctcagcgagagctg 1821
Db 1854 ccccccaccccaagctcagctgtgagagtgagagccacagactctcagcgagagctg 1913
QY 1822 actgacctgtgaagtgagagccatcagagagagagaggttgcccttctcttaaaagggcc 1881
Db 1914 actgacctgtgaagtgagagccatcagagagagagaggttgcccttctcttaaaagggcc 1973
QY 1882 tctgtgctgggagaaatccctcagagatcccaataaagtgtaaggtgtgagagggagccaa 1941
Db 1974 tctgtgctgggagaaatccctcagagatcccaataaagtgtaaggtgtgagagggagccaa 2033
QY 1942 gggaccaaagatagggcactgtgggtctatgtccacatacccaagttgttcgctcctg 2001
Db 2034 gggaccaaagatagggcactgtgggtctatgtccacatacccaagttgttcgctcctg 2093
QY 2002 agtcttctatgtgacacttaagtctgtctcctcagactcccaactgcttccctcctc 2061
Db 2094 agtcttctatgtgacacttaagtctgtctcctcagactcccaactgcttccctcctc 2153
QY 2062 ttccagagctgcttggggctcaagggcctgtactcatctggcaggtgacatgagatctg 2121
Db 2154 ttccagagctgcttggggctcaagggcctgtactcatctggcaggtgacatgagatctg 2213
QY 2122 ggaagctctctagagagatgagagccagagagctgtccacaaatgtcagaagcttgagat 2181
Db 2214 ggaagctctctagagagatgagagccagagagctgtccacaaatgtcagaagcttgagat 2273
QY 2182 gacactatctggcagacatcatctgtctctgtcatcatcttgaaacatattataagca 2241
Db 2274 gacactatctggcagacatcatctgtctctgtcatcatcttgaaacatattataagca 2333
QY 2242 ctgataataggtagcctgtgtgggtgatacagacattgactcaagatataagctcagact 2301
Db 2334 ctgataataggtagcctgtgtgggtgatacagacattgactcaagatataagctcagact 2393
QY 2302 cacagaattatagttataaaaaaacaagaagaacaacaatttgatcaaaagagaa 2361
Db 2394 cacagaattatagttataaaaaaacaagaagaacaacaatttgatcaaaagag-a 2452
|||||


```

Db 911 ctgcagctgaggaggagatgagctgtctgaaactaacaaccccgacagctgagc 970
QY 976 gggaaagagatcttctccctgtgtcccaactgtgtgacatgtacactaatatgttcaaa 1035
Db 971 gggaaagagatcttctccctgtgtcccaactgtgtgacatgtacactaatatgttcaaa 1030
QY 1036 ggcatacatagcttgcacaagaatcatctctactctcagagagacttgcacatcgagagacag 1095
Db 1031 ggcatacatagcttgcacaagaatcatctctactctcagagagacttgcacatcgagagacag 1090
QY 1096 atctccctgtcgaagaggcgagcttgcagagctgtgtcaactgtagaattcaacaagctgttc 1155
Db 1091 atctccctgtcgaagaggcgagcttgcagagctgtgtcaactgtagaattcaacaagctgttc 1150
QY 1156 aacgcggagagactggaaccccgaggaggtgtgagccgctgtctactgtgttggaagacacttca 1215
Db 1151 aacgcggagagactggaaccccgaggaggtgtgagccgctgtctactgtgttggaagacacttca 1210
QY 1216 ggtggtcttcacagacacttctactatgagacacatgtctgaaattccataatgtctgaaagag 1275
Db 1211 ggtggtcttcacagacacttctactatgagagcccatgtctgaaattccataatgtctgaaagag 1270
QY 1276 ctgcagctgtcatgtagagagagatgtatgtgtcatgtacagagccatctccctcttccccaagc 1335
Db 1271 ctgcagctgtcatgtagagagagatgtatgtgtcatgtacagagccatctccctcttccccaagc 1330
QY 1336 cgcacagaggtgtctgtcagacacgcgtgtgtgtgagccagctgtgcagagacaaattgcacattct 1395
Db 1331 cgcacagaggtgtctgtcagacacgcgtgtgtgtgagccagctgtgcagagacaaattgcacattct 1390
QY 1396 cggaaagctccataatgtgaatgtcaatctgcgcacccagcctgtctcatagattcttctgtccctgag 1455
Db 1391 cggaaagctccataatgtgaatgtcaatctgcgcacccagcctgtctcatagattcttctgtccctgag 1450
QY 1456 atcatgtgctatgtctcaacagagctctccgcagcatcaatgtctcagacacaccacagcgtctctg 1515
Db 1451 atcatgtgctatgtctcaacagagctctccgcagcatcaatgtctcagacacaccacagcgtctctg 1510
QY 1516 cgcataccaaagatacaaccccttctgtatgagccctcatagtcagagagatttctgcagactcaaa 1575
Db 1511 cgcataccaaagatacaaccccttctgtatgagccctcatagtcagagagatttctgcagactcaaa 1570
QY 1576 ggtagctgagcggcgtgccccttgtgtgtgacacacccctcgagagagcagccagaccccgacccctc 1635
Db 1571 ggtagctgagcggcgtgccccttgtgtgtgacacacccctcgagagagcagccagaccccgacccctc 1630
QY 1636 tgaagccgcacatcccgaggccaagaagatgtgaacactgtgcacaagccgcgaatgtccctgtct 1695
Db 1631 tgaagccgcacatcccgaggccaagaagatgtgaacactgtgcacaagccgcgaatgtccctgtct 1690
QY 1696 ggcctgtctcccttaaggaaattctctctatgagacgtgtgttagcatctctcagagaaagaca 1755
Db 1691 ggcctgtctcccttaaggaaattctctctatgagacgtgtgttagcatctctcagagaaagaca 1750
QY 1756 tgggtgtccccccacaccccgacttcaagtcctgtatgagagctgtgtagcatctctcagagaaagaca 1815
Db 1751 tgggtgtccccccacaccccgacttcaagtcctgtatgagagctgtgtagcatctctcagagaaagaca 1810
QY 1816 gagtgtcaactgtacctgtatgtcaagagacatcagagagaggaagtgtcccttctctctttaa 1875
Db 1811 gagtgtcaactgtacctgtatgtcaagagacatcagagagaggaagtgtcccttctctctttaa 1870
QY 1876 agggccctgtgtctgtgagggaatccctcctagatcccaactaaagtgtcaaggtgtgaaagg 1935
Db 1871 agggccctgtgtctgtgagggaatccctcctagatcccaactaaagtgtcaaggtgtgaaagg 1930
QY 1936 gaaccaagcgaacaagagatagcgcacatctgtgtgtatgtcccaataccacagtttgttcg 1995
Db 1931 gaaccaagcgaacaagagatagcgcacatctgtgtgtatgtcccaataccacagtttgttcg 1990
QY 1996 ttccctgtagctcttcaactgtactaacttaataagctgtgtctccacttccactgtgtcc 2055
Db 1991 ttccctgtagctcttcaactgtactaacttaataagctgtgtctccacttccactgtgtcc 2050

```

```

QY 2056 ctccctctccagagctgtctgtgtgtgtcgaagcctgtactcatcgcgaagtgtcatgtat 2115
Db 2051 ctccctctccagagctgtctgtgtgtgtgtcgaagcctgtactcatcgcgaagtgtcatgtat 2110
QY 2116 tctgtgtgagctctctcagaagaatgagaagcgaagcgcctgcacacaaatgtcagaagct 2175
Db 2111 tctgtgtgagctctctcagaagaatgagaagcgaagcgcctgcacacaaatgtcagaagct 2170
QY 2176 tggcatgacccatctccgcgcacatcatctgtgtctgtcgcacatcttggaacacattat 2235
Db 2171 tggcatgacccatctccgcgcacatcatctgtgtctgtcgcacatcttggaacacattat 2230
QY 2236 taagcaactgtataatgagtaagcctgtctgtgtgtgtatatacagatgtacatataagatcc 2295
Db 2231 taagcaactgtataatgagtaagcctgtctgtgtgtgtatatacagatgtacatataagatcc 2290
QY 2296 tgaagctcacagagtttatagttataaaaaaacaagaacaacaacaatttggatcaaaa 2355
Db 2291 tgaagctcacagagtttatagttataaaaaaacaagaacaacaacaatttggatcaaaa 2350
QY 2356 gggagaaatgtatagtgagacaaaagcagccagagaaatttccctgtgtgtgtatgtatgtc 2415
Db 2351 gggag-aaatgtatagtgagacaaaagcagccagagaaatttccctgtgtgtgtatgtatgtc 2409
QY 2416 tgaatgtcagagcactgtgtgtgtacccaaggtgtcccgagagacatgtatgtatgtatgtc 2475
Db 2410 tgaatgtcagagcactgtgtgtgtgtacccaaggtgtcccgagagacatgtatgtatgtatgtc 2469
QY 2476 gggacacaacactgtcagcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2535
Db 2470 gggacacaacactgtcagcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2529
QY 2536 acttgtatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2595
Db 2530 acttgtatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2589
QY 2596 ctacgctgtacacatcagtttaaacacacccggagagaaacattacatgtacacttaatttc 2655
Db 2590 ctacgctgtacacatcagtttaaacacacccggagagaaacattacatgtacacttaatttc 2649
QY 2656 tgtgtacacacacttatctcacaagcctaagaaggtatgaagtgccgtcctgtttatagcca 2715
Db 2650 tgtgtacacacacttatctcacaagcctaagaaggtatgaagtgccgtcctgtttatagcca 2709
QY 2716 ctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2775
Db 2710 ctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2769
QY 2776 cctaagacactagtttctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2835
Db 2770 cctaagacactagtttctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2829
QY 2836 attaaataatgtacttttggctaaa 2860
Db 2830 attaaataatgtacttttggctaaa 2854

```

RESULT 4
 X59974 standard; DNA; 3243 BP.
 XX
 AC X59974;
 DT 04-AUG-1999 (first entry)
 XX
 DE DNA encoding an intranuclear receptor protein.
 KW Human; intranuclear receptor protein; drug development; diagnosis;
 KW treatment; ss.
 XX
 OS Homo sapiens.
 XX

QY 1628 gagccctctgagccgacactcccgccggaagaacagatgacatgcacagagccgacaat 1687
 DB 1809 gagccctctgagccgacactcccgccggaagaacagatgacatgcacagagccgacaat 1868
 QY 1688 ggcctctgagccctgtctccctagaggaattcctgtctatgacagctgtgtagcatctctag 1747
 DB 1869 ggcctctgagccctgtctccctagaggaattcctgtctatgacagctgtgtagcatctctag 1928
 QY 1748 gaagagacatgggtgcccccccccaagttcaagttctgtatggaggtgaagccacaactct 1807
 DB 1929 gaagagacatgggtgcccccccccaagttcaagttctgtatggaggtgaagccacaactct 1988
 QY 1808 tacgtgagagatgacatgacatgcgtatgagatgacatgcacagaggaaggttgcctcttc 1867
 DB 1989 tacgtgagagatgacatgacatgcgtatgagatgacatgcacagaggaaggttgcctcttc 2048
 QY 1868 ctttaaaaggccctgtgtgtctgtggaagaatccctcagatcccaataaagtgtcaaggt 1927
 DB 2049 ctttaaaaggccctgtgtgtctgtggaagaatccctcagatcccaataaagtgtcaaggt 2108
 QY 1928 ggggaagggacacagccagcaagagatagggcactgtgggtctatgtcccaatacccaagt 1987
 DB 2109 ggggaagggacacagccagcaagagatagggcactgtgggtctatgtcccaatacccaagt 2168
 QY 1988 ttgttcgctccctgagctcttccatgtgacactcctaatagtctgtctcccaactccac 2047
 DB 2169 ttgttcgctccctgagctcttccatgtgacactcctaatagtctgtctcccaactccac 2228
 QY 2048 tcgttcctccctctctccgagctgtctgtgtggtcacaaggccctgtacatccatccgaggtg 2107
 DB 2229 tcgttcctccctctctccgagctgtctgtgtggtcacaaggccctgtacatccatccgaggtg 2288
 QY 2108 catgtgatactgtgtggaggtccctctagagagatgagaagccagagagctgtgacccaagt 2167
 DB 2289 catgtgatactgtgtggaggtccctctagagagatgagaagccagagagctgtgacccaagt 2348
 QY 2168 cagaagctgtgacatgcacatccctccgacacatcactctgtgtctgcacatcatttga 2227
 DB 2349 cagaagctgtgacatgcacatccctccgacacatcactctgtgtctgcacatcatttga 2408
 QY 2228 cacatataaagcactgataaataagtaggcctgtgtgtggtatatacagattgactcagat 2287
 DB 2409 cacatataaagcactgataaataagtaggcctgtgtgtggtatatacagattgactcagat 2468
 QY 2288 atagatcctgagctacacagagtttatagttataaaaaaacacagaaacacaaacatttg 2347
 DB 2469 atagatcctgagctacacagagtttatagttataaaaaaacacagaaacacaaacatttg 2528
 QY 2348 gatcaaaaggagaataatgatagtacaaaagcagacagaagattccctgtgtgaggtg 2407
 DB 2529 gatcaaaaggag-aaatgataatgataaaaagcagacagaagattccctgtgtgaggtg 2587
 QY 2408 ctgagctgtgatagtgacagcactgtgtgtacccaagtgaaggtcccgagagacatgagctgt 2467
 DB 2588 ctgagctgtgatagtgacagcactgtgtgtacccaagtgaaggtcccgagagacatgagctgt 2647
 QY 2468 aggaagaaggagacaaactgcagcactgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2527
 DB 2648 aggaagaaggagacaaactgcagcactgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2707
 QY 2528 tctttgcccactgtgaggggcccgt 2587
 DB 2708 tctttgcccactgtgaggggcccgt 2767
 QY 2588 tgacaaggctacgctgtgacaatcagttaaacacacggaagaacacattatcatgacact 2647
 DB 2768 tgacaaggctacgctgtgacaatcagttaaacacacggaagaacacattatcatgacact 2827
 QY 2648 tatattctgtgtacacatctattctcaaaagctaaaggtatgaaggtgctgtcctgtgt 2707
 DB 2828 tatattctgtgtacacatctattctcaaaagctaaaggtatgaaggtgctgtcctgtgt 2887

QY 2708 tatagccactgtgagtaaaatttttttgcattttcacaaattatattatataagc 2767
 DB 2888 tatagccactgtgagtaaaatttttttgcattttcacaaattatattatataagc 2947
 QY 2768 attcacacctaaagacatgttttgggaaatgtagccctgggttttaattgcaataaag 2827
 DB 2948 attcacacctaaagacatgttttgggaaatgtagccctgggttttaattgcaataaag 3007
 QY 2828 caaaagacatataatgacttttggcttaa 2860
 DB 3008 caaaagacatataatgacttttggcttaa 3040

RESULT 5

ID X56243 standard; cDNA; 2802 BP.
 AC X56243;
 DT 16-JUL-1999 (first entry)
 XX

DE Human vitamin D-receptor related gamma 2 protein encoding cDNA.

KW Human; vitamin D receptor related protein; VDR; obesity; diabetes;
 anorexia; rheumatoid arthritis; lipoprotein defect; hyperlipidaemia;
 hypercholesterolaemia; hyperlipoproteinaemia; osteoporosis; tumour;
 hyperproliferative skin disorder; hyperthyroidism; ss.

OS Homo sapiens.

PN WO919354-A1.

PD 22-APR-1999.

PF 31-APR-1998; 98MO-SE01548.

PR 31-MAR-1998; 98SE-0001148.

PR 14-OCT-1997; 97SE-0003745.

PA (PHMA) PHARMACIA & UPJOHN AB.

PI Berkenstam A, Dahlberg M;

PT WPI: 1999-302508/25.

DR P-P-SDB; Y09516.

XX New vitamin D receptor related (VDR) polypeptides, useful for

XX treating obesity, diabetes, anorexia and rheumatoid arthritis

XX Claim 2; Page 21-22; 35pp; English.

CC The present sequence encodes a human vitamin D receptor related (VDR)
 CC polypeptide. Human VDR polypeptides and substances which affect VDR
 CC signal transduction, can be used for treating metabolic, proliferative
 CC or inflammatory conditions. They can be used in the manufacture of a
 CC medicament for treating the following conditions: obesity, diabetes,
 CC anorexia, lipoprotein defects, hyperlipidaemia, hypercholesterolaemia or
 CC hyperlipoproteinaemia and osteoporosis, rheumatoid arthritis, benign and
 CC malignant tumours, hyperproliferative skin disorders or hyperthyroidism.
 CC Nucleic acid vectors encoding for expression of a VDR polypeptide can
 CC be used for treating metabolic, proliferative or inflammatory conditions,
 CC by introducing them into a mammal. The introduced nucleic acid is then
 CC capable of transforming a cell in vivo and then polypeptide is expressed.
 CC A substance affecting VDR signal transduction, such as an agonist or
 CC antagonist can be used for the manufacture of a medicament for treating
 CC metabolic, proliferative or inflammatory condition.
 CC N.B. The specification specifically claims the VDR nucleic acid and
 CC polypeptide sequences given in figures 1, 4, 7 and 8, but no figures
 CC are given in the specification.

SQ Sequence 2802 BP; 723 A; 715 C; 755 G; 609 T; 0 other;


```

Db 1720 ggcctctccctaggaattctctctatagacagctgctagcattctctcagaagaca 1779
Qy 1756 tgggtgccccccacccacgctcagctctgtagggagtagaagccacagctctttagtga 1815
Db 1780 tgggtgccccccacccacgctcagctctgtagggagtagaagccacagctctttagtga 1839
Qy 1816 ggtgctacagctctgtagttagcagacatcagaagagcaaggttgccttctctttaa 1875
Db 1840 ggtgctacagctctgtagttagcagacatcagaagagcaaggttgccttctctttaa 1899
Qy 1876 aggcctctgctgctgtagggagaaatccctcagaatcccaaaagtctcaagtgtagaag 1935
Db 1900 aggcctctgctgctgtagggagaaatccctcagaatcccaaaagtctcaagtgtagaag 1959
Qy 1936 gaccagagcacaagaatagagcctctggtgctatgcccacataccagtttctgc 1995
Db 1960 gaccagagcacaagaatagagcctctggtgctatgcccacataccagtttctgc 2019
Qy 1996 ttcctgagctcttctatgctacccataagctctgctcctccactccactgctccc 2055
Db 2020 ttcctgagctcttctatgctacccataagctctgctcctccactccactgctccc 2079
Qy 2056 cctctcttcagagctgcttctgtagcctcaagagcctgtactcagcagagtgatgagta 2115
Db 2080 cctctcttcagagctgcttctgtagcctcaagagcctgtactcagcagagtgatgagta 2139
Qy 2116 tctgtgg 2122
Db 2140 tctgtgg 2146

```

RESULT 8

X89080
ID X89080 standard; DNA; 2068 BP.

XX X89080;

DT 14-SEP-1999 (first entry)

DE Human steroid and xenobiotic receptor (SXR) encoding DNA.

XX Nuclear receptor; SXR; steroid and xenobiotic receptor; RXR; human;
 KW retinoid X receptor; P450 gene; steroid hormone; steroid metabolism;
 KW phytoestrogen; calcium-channel blocker; steroid toxicity; tuberculosis;
 KW breast cancer; osteoporosis; Cushing syndrome; vitilism; hirsutism;
 KW polycystic ovarian disease; cancer; colorectal; prostatic; ss.

XX Homo sapiens.

XX Key Location/Qualifiers

FT CDS 583..1887

FT FT /*tag= a /transl_except= (pos:1261..1263, aa:Xaa)

FT FT /note= "Xaa= unknown; the start codon is not indicated"

XX W09935246-A1.

XX 15-JUL-1999.

XX 08-JAN-1999; 99WO-US00490.

XX 09-JAN-1998; 98US-0005286.

XX (SALK) SALK INST BIOLOGICAL STUDIES.

XX Blumberg B, Evans RM;

XX WPI; 1999-419349/35.

XX P-PSDB; Y21799.
 DR New steroid and xenobiotic receptor, used to identify modulators for
 PT controlling metabolism of steroids and xenobiotics, e.g. reducing

PT their toxicity
 XX
 PS Claim 10; Fig 1A; 83pp; English.

CC The invention relates to a novel nuclear receptor polypeptide, designated
 CC SXR (steroid and xenobiotic receptor). SXR (1) forms a heterodimer with
 CC retinoid X receptor (RXR), (11) binds to a direct or inverted repeat
 CC response element motif based on the half-site AGTCA, (11) activates
 CC transcription through response elements present in steroid-inducible P450
 CC genes, in response to a wide variety of natural and synthetic steroid
 CC hormones and (1V) is prominently expressed in liver and intestine. SXR
 CC regulates expression of catabolic enzymes, in response to many different
 CC steroids, and thus affects metabolism. SXR is a broad specificity, low-
 CC affinity receptor for reducing excessive levels of steroids in the
 CC circulation. (Ant)agonists of SXR are used to regulate metabolism of
 CC steroids particularly phytoestrogens or calcium-channel blockers, to
 CC reduce steroid toxicity in subjects being treated with steroids, e.g. in
 CC cases of tuberculosis (treated with rifampin and related compounds),
 CC breast cancer (treated with tamoxifen, raloxifen etc.) or osteoporosis
 CC (treated with Vitamin K), or to slow metabolism of therapeutic steroids.
 CC Also, modulating endogenous SXR is used to treat disease, particularly
 CC an agonist is used where endogenous steroid levels are excessive (e.g.
 CC Cushing syndrome; vitilism and hirsutism in women; polycystic ovarian
 CC disease; 11 beta-, 17- or 21-hydroxylase deficiency; 3 beta-hydroxysteroid
 CC dehydrogenase deficiency, or breast, colorectal or prostatic cancer),
 CC while antagonists are used where endogenous steroid levels are too low.
 CC Cells that express SXR are used to identify compounds likely to be
 CC involved in undesirable drug interactions. Antibodies specific for SXR
 CC are used in immunohistochemical testing for studying distribution/
 CC expression density of SXR, also for diagnosis and therapeutically as
 CC antagonist. The present sequence represents the longest SXR cDNA clone
 CC encoding the SXR polypeptide.

CC Sequence 2068 BP; 520 A; 541 C; 586 G; 420 T; 1 other;

Query Match 58.0%; Score 1686.2; DB 20; Length 2068;

Best Local Similarity 98.7%; Pred. No. 0;

Matches 1730; Conservative 0; Mismatches 19; Indels 3; Gaps 3;

```

Qy 16 gaatcagatgtaattcgtggaacgggaaggaagcactgcttacttacttgggaat 75
Db 319 gaatacatacctatgactagcagggagaaggaagcactgcttacttacttgggaat 378
Qy 76 ctggcctcagctctggaagcagagtggttcaacgtgtaaaagaagaagataagcaata 135
Db 379 ctggcctcagctctggaagcagagtggttcaacgtgtaaaagaagaagataagcaata 438
Qy 136 ctccctgctgaacaaagcagcgcccttggtgaagctactccttgatcggatccttgc 195
Db 439 ctccctgctgaacaaagcagcgcccttggtgaagctactccttgatcggatccttgc 498
Qy 196 accggaattgticaaagtggaccccaagggaagagtcgagcaaaagacttacccaagc 255
Db 499 accggaattgticaaagtggaccccaagggaagagtcgagcaaaagacttacccaagc 558
Qy 256 agtccaagaagcccaagaagaacactggaggtgagacccaagaagaactggaaactgct 315
Db 559 agtccaagaagcccaagaagaacactggaggtgagacccaagaagaactggaaactgct 618
Qy 316 gacttgaactgtgagacacagagctggttcctcggaagacccaggttcaacgagat 375
Db 619 gacttgaactgtgagacacagagctggttcctcggaagacccaggttcaacgagat 678
Qy 376 gaggaaagtcggaagtcaccaaatctgcgtgtatgtgggacaagccctggctatcac 435
Db 679 gaggaaagtcggaagtcaccaaatctgcgtgtatgtgggacaagccctggctatcac 738
Qy 436 ttaactgtcatgcatgtgaaagatcaagaggttttcaagagggccatgaagcaac 495
Db 739 ttaactgtcatgcatgtgaaagatcaagaggttttcaagagggccatgaagcaac 798
Qy 496 gcccggtgagggtgccccttcggaagggcgctgtagagatcacccggaaagcccgagca 555

```

```

|||||
Db 799 gccgcgtgagtgcccccttcaggaaaggcgccctgagatcaccccggaagaccgcgcga 858
QY 556 cagtgccagagccctgcgccttcgcaagtgcttgagagacggcagatgaagagatgctc 615
Db 859 cagtgccagagccctgcgccttcgcaagtgcttgagagacggcagatgaagagatgctc 918
QY 616 atgtccagagagccctgcgccttcgcaagtgcttgagagacggcagatgaagagatgctc 675
Db 919 atgtccagagagccctgcgccttcgcaagtgcttgagagacggcagatgaagagatgctc 978
QY 676 acaaggagacacagcagctgagagtgcaagagtgcaagagagcagcagatgagatcag 735
Db 979 acaaggagacacagcagctgagagtgcaagagtgcaagagagcagcagatgagatcag 1038
QY 736 gagctgagagagcagctgagagtgcaagagtgcaagagtgcaagagtgcaagagtgca 795
Db 1039 gagctgagagagcagctgagagtgcaagagtgcaagagtgcaagagtgcaagagtgca 1098
QY 796 cggctgcagagagtgctgagagtgcaagagtgcaagagtgcaagagtgcaagagtgca 855
Db 1099 cggctgcagagagtgctgagagtgcaagagtgcaagagtgcaagagtgcaagagtgca 1158
QY 856 agggagagagcagcagctgagagtgcaagagtgcaagagtgcaagagtgcaagagtgca 915
Db 1159 agggagagagcagcagctgagagtgcaagagtgcaagagtgcaagagtgcaagagtgca 1218
QY 916 ctgc-cagctgcgggggagagtgcaagtgctgagagtgcaagagtgcaagagtgcaag 974
Db 1219 ctgc-cagctgcgggggagagtgcaagtgctgagagtgcaagagtgcaagagtgcaag 1278
QY 975 cgggagagagagctctccctgcgccttcgcaagtgctgagagtgcaagagtgcaagagtg 1034
Db 1279 c-ggagagagagctctccctgcgccttcgcaagtgctgagagtgcaagagtgcaagagtg 1337
QY 1035 agggagagagcagcagctgagagtgcaagagtgcaagagtgcaagagtgcaagagtgca 1094
Db 1338 agggagagagcagcagctgagagtgcaagagtgcaagagtgcaagagtgcaagagtgca 1397
QY 1095 gatctccctgcgggggagagtgcaagtgctgagagtgcaagagtgcaagagtgcaag 1154
Db 1398 gatctccctgcgggggagagtgcaagtgctgagagtgcaagagtgcaagagtgcaag 1457
QY 1155 caacgcggagagcagcagctgagagtgcaagtgctgagagtgcaagagtgcaagagtgca 1214
Db 1458 caacgcggagagcagcagctgagagtgcaagtgctgagagtgcaagagtgcaagagtgca 1517
QY 1215 aggtgagctcagagagcagcagctgagagtgcaagtgctgagagtgcaagagtgcaag 1274
Db 1518 aggtgagctcagagagcagcagctgagagtgcaagtgctgagagtgcaagagtgcaag 1577
QY 1275 gctgcagctcagagagagtgagtgctgagagtgcaagtgctcctctctccagag 1334
Db 1578 gctgcagctcagagagagtgagtgctgagagtgcaagtgctcctctctccagag 1637
QY 1335 ccgcgcagagtgctgagagcagcagctgagagtgcaagtgctgagagtgcaagtgctgag 1394
Db 1638 ccgcgcagagtgctgagagcagcagctgagagtgcaagtgctgagagtgcaagtgctgag 1697
QY 1395 tctgagagctcagagtgagtgcaagtgctgagagtgcaagtgctgagagtgcaagtg 1454
Db 1698 tctgagagctcagagtgagtgcaagtgctgagagtgcaagtgctgagagtgcaagtg 1757
QY 1455 gatcagtgctatgctacagagtgctgagagtgcaagtgctgagagtgcaagtgctgag 1514
Db 1758 gatcagtgctatgctacagagtgctgagagtgcaagtgctgagagtgcaagtgctgag 1817
QY 1515 gcgcagcagagagagagagagagagagagagagagagagagagagagagagagagagag 1574
Db 1818 gcgcagcagagagagagagagagagagagagagagagagagagagagagagagagagag 1877
QY 1575 aggtgagcagagagagagagagagagagagagagagagagagagagagagagagagagag 1634
|||||
Db 1878 aggtgagcagagagagagagagagagagagagagagagagagagagagagagagag 1936
QY 1635 ctgagcagcagagagagagagagagagagagagagagagagagagagagagagagag 1694
Db 1937 ctgagcagcagagagagagagagagagagagagagagagagagagagagagagagag 1996
QY 1695 tggcagcagcagagagagagagagagagagagagagagagagagagagagagagag 1754
Db 1997 tggcagcagcagagagagagagagagagagagagagagagagagagagagagagag 2056
QY 1755 atgggtgcccc 1766
Db 2057 atgggtgcccc 2068

RESULT 9
ID X59967
XX X59967 standard; DNA: 1374 BP.
AC X59967:
XX
XX
XX 04-AUG-1999 (first entry)
DT
XX
XX
XX DNA encoding an intranuclear receptor protein.
DE
XX Human; intranuclear receptor protein; drug development; diagnosis;
KW treatment; ss.
XX
XX Homo sapiens.
XX
XX JP1127872-A.
XX
XX 18-MAY-1999.
PD
XX
XX 07-AUG-1998; 98UP-0224172.
PF
XX
XX 11-AUG-1997; 97UP-0230335.
PR
XX
XX (NISB) JAPAN TOBACCO INC.
PA
XX
XX WPI: 1999-350330/30.
XX
XX P-PSDB; Y15932.
DR
XX
XX The present sequence encodes a human intranuclear receptor protein.
XX CC The nucleic acid sequence was isolated from a human adult cDNA
XX CC library using a swifflish ANO23 derived probe. The protein can
XX CC be used for the development of drugs and diagnosis and treatment
XX CC of various diseases.
XX
XX
XX Sequence 1374 BP: 329 A; 375 C; 406 G; 264 T; 0 other;
SQ

Query Match 47.3%; Score 1374; DB 20; Length 1374;
Best Local Similarity 100.0%; Pred. No. 6e-272;
Matches 1374; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 211 gtggagccagggagagagagagagagagagagagagagagagagagagagagagagag 270
Db 1 gtggagccagggagagagagagagagagagagagagagagagagagagagagagagag 60
QY 271 gaagcaaaccttgagagtgagagagagagagagagagagagagagagagagagagagag 330
Db 61 gaagcaaaccttgagagtgagagagagagagagagagagagagagagagagagagagag 120
QY 331 gaagcaaaccttgagagtgagagagagagagagagagagagagagagagagagagagag 390
Db 121 gaagcaaaccttgagagtgagagagagagagagagagagagagagagagagagagagag 180

```



```

QY 542 ggaagaccgagacagtcagcctgcgcctcgcgaagtgccttgagagacgcagca 601
DB 380 ggaagaccgagacagtcagcctgcgcctcgcgaagtgccttgagagacgcagca 439
QY 602 agaagagagatgacatgctcgcagagagccgtgagagagagagagagagagagag 661
DB 440 agaagagagatgacatgctcgcagagagccgtgagagagagagagagagagagag 499
QY 662 aaaaaagtgaaagagagagagagagagagagagagagagagagagagagagagag 721
DB 500 aaaaaagtgaaagagagagagagagagagagagagagagagagagagagagagag 559
QY 722 ggaatgacagagagagagagagagagagagagagagagagagagagagagagag 781
DB 560 ggaatgacagagagagagagagagagagagagagagagagagagagagagagag 619
QY 782 attcaagaattccgagctgcagagagtgctagcagtgctgcagagtgccagagctc 841
DB 620 attcaagaattccgagctgcagagagtgctagcagtgctgcagagtgccagagctc 679
QY 842 tgcagagcccatcgcagagagagagagagagagagagagagagagagagagagag 901
DB 680 tgcagagcccatcgcagagagagagagagagagagagagagagagagagagagag 739
QY 902 ctttgaagctctctgcagctgcgagagagagagagagagagagagagagagagag 961
DB 740 ctttgaagctctctgcagctgcgagagagagagagagagagagagagagagagag 799
QY 962 cagcgcagagtgagcgcgagagagagagagagagagagagagagagagagagag 1021
DB 800 cagcgcagagtgagcgcgagagagagagagagagagagagagagagagagagag 859
QY 1022 cctacatgctcaaaagagacatcagcttgcgaaagtatctctcacttccagagag 1081
DB 860 cctacatgctcaaaagagacatcagcttgcgaaagtatctctcacttccagagag 919
QY 1082 ccatcgaagagacagatctcctcgcagagagagagagagagagagagagagagag 1141
DB 920 ccatcgaagagacagatctcctcgcagagagagagagagagagagagagagagag 979
QY 1142 tcaacacagtgctcaacgcgcgagagagagagagagagagagagagagagagag 1201
DB 980 tcaacacagtgctcaacgcgcgagagagagagagagagagagagagagagagag 1039
QY 1202 tggagagacatgagtggtctccagcaactctactgagagccatctgctgaattcc 1261
DB 1040 tggagagacatgagtggtctccagcaactctactgagagccatctgctgaattcc 1099
QY 1262 acatgctgaagagagctgcagctgcagtgagagagagagagagagagagagagag 1321
DB 1100 acatgctgaagagagctgcagctgcagtgagagagagagagagagagagagagag 1159
QY 1322 tcttctcccaagaccgcagagtgctgcagcagccgctggtgagcagagctgcagag 1381
DB 1160 tcttctcccaagaccgcagagtgctgcagcagccgctggtgagcagagctgcagag 1219
QY 1382 aattgcacattacttgagtgctcattgaattgaattgaattgaattgaattgaatt 1441
DB 1220 aattgcacattacttgagtgctcattgaattgaattgaattgaattgaattgaatt 1279
QY 1442 tcttgcctgaagatcatgagctatgctacacgagagctccgcagacatctcagaca 1501
DB 1280 tcttgcctgaagatcatgagctatgctacacgagagctccgcagacatctcagaca 1339
QY 1502 cccagcgcgtgctgcagcatcagagacatacacccttgcctacgcctcctcagagag 1561
DB 1340 cccagcgcgtgctgcagcatcagagacatacacccttgcctacgcctcctcagagag 1399
QY 1562 tgttcgcatcagagtgagctga 1584
DB 1400 tgttcgcatcagagtgagctga 1422

```

```

RESULT 11
ID X59966 standard; DNA; 1305 BP.
XX
XX AC X59966;
XX
XX 04-AUG-1999 (first entry)
XX
DE DNA encoding an intranuclear receptor protein.
XX
XX Human; intranuclear receptor protein; drug development; diagnosis;
XX treatment; ss.
XX
XX Homo sapiens.
XX
XX JP11127872-A.
XX
XX 18-MAY-1999.
XX
XX 07-AUG-1998; 98JP-0224172.
XX
XX 11-AUG-1997; 97JP-0230335.
XX
XX (NIBS ) JAPAN TOBACCO INC.
XX
XX WPI: 1999-350330/30.
XX
XX P-P-SDB; Y15931.
XX
XX New intranuclear receptor protein - useful for drug development and
XX diagnosis and treatment of disease
XX
XX PS Claim 4; Page 22-23; 38pp; Japanese.
XX
XX CC The present sequence encodes a human intranuclear receptor protein.
XX CC The nucleic acid sequence was isolated from a human adult cDNA
XX CC library using a swellfish ANO23 derived probe. The protein can
XX CC be used for the development of drugs and diagnosis and treatment
XX CC of various diseases.
XX
XX SQ Sequence 1305 BP; 304 A; 356 C; 386 G; 259 T; 0 other;

Query Match 44.9%; Score 1305; DB 20; Length 1305;
Best Local Similarity 100.0%; Pred. No. 7.5e-258;
Matches 1305; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 280 ctgagagtgagaccacaaagagagagagagagagagagagagagagagagagagag 339
DB 1 ctgagagtgagaccacaaagagagagagagagagagagagagagagagagagagagag 60
QY 340 gaagtcgttctctgagagagagagagagagagagagagagagagagagagagagag 399
DB 61 gaagtcgttctctgagagagagagagagagagagagagagagagagagagagagag 120
QY 400 tgcgcgtatgctgagagagagagagagagagagagagagagagagagagagagagag 459
DB 121 tgcgcgtatgctgagagagagagagagagagagagagagagagagagagagagagag 180
QY 460 tgaagagagcttcttcagagagagagagagagagagagagagagagagagagagag 519
DB 181 tgaagagagcttcttcagagagagagagagagagagagagagagagagagagagag 240
QY 520 aagagcgcctgcagagatcaccgagagagagagagagagagagagagagagagagag 579
DB 241 aagagcgcctgcagagatcaccgagagagagagagagagagagagagagagagagag 300
QY 580 aagtgccctgagagagagagagagagagagagagagagagagagagagagagagag 639
DB 301 aagtgccctgagagagagagagagagagagagagagagagagagagagagagagag 360
QY 640 aggcgcgccttgatcaagcggagagagagagagagagagagagagagagagagagag 699

```

[illegible]

KM Human: Intracellular receptor protein; drug development; diagnosis;
 KW treatment, ss.
 XX Homo sapiens.
 OS
 PN JF1127872-A.
 XX
 PD 18-MAY-1999.
 PF 07-AUG-1998; 98JP-0224172.
 PR 11-AUG-1997; 97JP-0230335.
 XX
 PA (NINSB) JAPAN TOBACCO INC.
 XX WPI; 1999-350330/30.
 DR
 PT New intranuclear receptor protein - useful for drug development and
 PT diagnosis and treatment of disease
 XX
 PS Disclosure; Page 32; 38pp; Japanese..
 CC The specification describes a human intranuclear receptor protein.
 CC The nucleic acid sequence was isolated from a human adult CDNA
 CC library using a swellfish ANO3 derived probe. The protein can
 CC be used for the development of drugs and diagnosis and treatment
 CC of various diseases. The present sequence was used in the course of
 CC the invention.

SQ Sequence 644 BF; 175 A; 161 C; 193 G; 115 T; 0 other;

```

Query Match          20.9%; Score 606.6; DB 20; Length 644;
Best Local Similarity 99.2%; Pred.No. 3,1e-115;
Matches 620; Conservative 0; Mismatches 4; Indels 1; Gaps 1.

QY 55 tgccttacttcgtggtgaatcctcgccctcgcaagcagaattttacacgtgaaan 114
DB 12 tgccttacttcagt-gaatcttgccctcagcctcgcaagtggttccacgttgaa 70
   |||||||
QY 115 aagaagaagataagctaatactctctgtaacaaggcagcgctcctgtgaaagct 174
DB 71 aagaagaagataagctaatactctctgtaacaaggcagcgctcctgtgaaagct 130
   |||||||
QY 175 actccttgatcgatcctcttgccaccggatgtgttcaaagtggccccagggagaaagtcgga 234
DB 131 actccttgatcgatcctcttgccaccggatgtgttcaaagtggccccagggagaaagtcgga 190
   |||||||
QY 235 gcaaaagaacttaccaccaagcagctccaagagggcccagaagaacacccgggggtgaagccc 284
DB 191 gcaaaagaacttaccaccaagcagctccaagagggcccagaagaacacccgggggtgaagccc 250
   |||||||
QY 295 aaagaagaagctgggaaccatgctgaacttgtgaacgtgtgagagcacagaaggtcgtgtccgga 354
DB 251 aaagaagaagctgggaaccatgctgaacttgtgaacgtgtgagagcacagaaggtcgtgtccgga 310
   |||||||
QY 355 aagcccgctgttcaacgcgacgataggagaagtcgggaaggtccccaatatctgcgtgtatgtgg 414
DB 311 aagcccgctgttcaacgcgacgataggagaagtcgggaaggtccccaatatctgcgtgtatgtgg 370
   |||||||
QY 415 gacaagaagcacttgcttatcaacttcaatgtatgacatgttgaagagatgcaagagctttttc 474
DB 371 gacaagaagcacttgcttatcaacttcaatgtatgacatgttgaagagatgcaagagctttttc 430
   |||||||
QY 475 aggaaggcgccatgtgaaaagcaaacgcccggtcgaggttcccccttcgggaagggcgctcgag 534
DB 431 aggaaggcgccatgtgaaaagcaaacgcccggtcgaggttcccccttcgggaagggcgctcgag 490
   |||||||
QY 535 atccaccggagaagcccgcgacagatgtgccagggcttcgcgctcgcaagatgtccttgagagagc 594
DB 491 atccaccggagaagcccgcgacagatgtgccagggcttcgcgctcgcaagatgtccttgagagagc 550
   |||||||
QY 595 ggcataagaagaagatgatacatgtctcgacagagggccgttggaggaagggcgggccttgatc 654
```


D	392	gtacacagctatagaggtatgtataatgtaaagagaaagagagagagaggtcttgaaagagcagt	451
O	691	cttgagatgtgaaaggtctgtacaaagagagacagcgatgatatgtatcagagagactgtatgtacgact	750
D	452	cttggg---gcccaagctatctctaaagaaacaaacagacacatacatatagacatctctgtgagcgcc	508
O	751	cagaatgaaaacccctttgacacatgactctctcccatcttcacaaagaaattctcgggtcgcaag-----	805
D	509	cacacacaaagacactatgtaccccccaacgctgtgactctcagggagactccgggtccctcagaattcgt	568
O	806	-gggtgcttgcagctgtgctctgcagatgtgcacagatctctcgcaggtcccatctgaaagagaaagaa	864
D	569	atggagacgaaagtacaaagagagactatctctccaaagcccaacacatcagactctctccggaaactcc	628
O	865	gtctgcacaggtgagagcagaggtccgcgaagaaatctgtggtctcttgaaaggtctctctgcgcagcg	924
D	629	tctctctccagctctctacccctgtacacacacatccatctgaagaaatagatgagaaacatccggtctt	688
O	925	cggagggagagatgtgcagtgcttctcgaacacaaacacccacagccgaacagtgtgcggagaaagag	984
D	689	tccaaacctgatactgaacagagagagatctctgatacccgctctgtgactctgcggccgctct	748
O	985	attcttccctctgtgcgccccacatgtgctctcaatgtctcaaacatcaatgltcaaaagagcatatc	1044
D	749	ccctctcccatgtctgcgccccacatctgtgtaacctgttcaagtatacagacatccaaagagctatcc	808
O	1045	agctttgcacaaagatcacctccacatctcggagagactgtgcacatctgcagagagccagatctcccg	1104
D	809	ggctttgcacaaagatcacctccacatctcggagagactgtgcacatctgcagagagccagatctcccg	868
O	1105	ctgaaagggggccggtcttcgagctgtgttcaaatgaaatcacacaaagtgltcaaaacggcgagag	1164
D	869	cttaaaagtaaaagagcattgagtgtatcagttaacgtctcaacaaacaggtctcttaccacatgagt	928
O	1165	acgggaacactggagagtggtgcgggtgtctctacacgtctgtgaagaaacagt-----ca	1215
D	929	gataatgtctcggagacgtgtgcagccagagacataaagaatacgaatacgcatacgcagatgtctcaaa	988
O	1216	gggtgccttcacaaacactctcaactctgagagcccaatgcgtctgaataatcacatacatgctgaagag	1275
D	989	gctcgggcacaaacccctgtgagctgtatcgaagccctctataaagtctcagagtggtgcgtcgaagag	1048
O	1276	ctgcagactgtcagaaagagagatgattgtctgatactgacagagcatctccctctctccacagac	1335
D	1049	ctgaataatctaaatgaaagaaagacatgtctctctctcaatgacatctgacatgtctctcccgac	1108
O	1336	cggccaggggtgtctctcagacacggcggtgtgagacacagctgaaagagaaatctgcatact	1395
D	1109	cggacctgggtgtccaaagacagcaagctgtgtgaaagacatcaatgaagacggctataatccaaacg	1168
O	1396	ctgaagctctacatgaaatgaaatgcgaatcggcccccagcgctctcataggtctctgtcttcgaag	1455
D	1169	ctgcagaaacctaatatccgtctgcgcacacacggcccccaagcgagcaacagatctatctgaag	1228
O	1456	ataatgctctatgtctacacggagctcgcagacataatgattctaaagacacccagc	1507
D	1229	atgataccagaacactgcgcgaactctgcgaagctctcaacagaggaacacatctccaaac	1280
RESULT	15		
ID	X34789		
AC	X34789		
XX			
XX	06-VUL-1999	(first entry)	
DE		Nucleotide sequence of human vitamin D receptor (VDR) gene transcript	9
XX		Vitamin D receptor; VDR; hVDR; variant; isoform; SRC-1; GRIP-1; TRIPB;	
KM		cofactor; human; ds.	
XX			
XX		Homo sapiens	

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 18, 2001, 13:43:40 ; Search time 6591.86 Seconds
(without alignments)
225.367 Million cell updates/sec

Title: US-09-143-828-1

Perfect score: 2905
Sequence: 1 cctctgaaggtctagaatc.....aaaaaaaaaaaaaaaaaaaa 2905

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 1118133 seqs, 2558875100 residues

Total number of hits satisfying chosen parameters: 2236266

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 08
Maximum Match 1008

Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba1:*
2: gb_ba2:*
3: gb_om:*
4: gb_ov:*
5: gb_ph:*
6: gb_pl1:*
7: gb_pl2:*
8: gb_pr1:*
9: gb_pr2:*
10: gb_pr3:*
11: gb_ro:*
12: gb_ay:*
13: gb_un:*
14: em_fun:*
15: em_hum1:*
16: em_hum2:*
17: em_in:*
18: em_om:*
19: em_or:*
20: em_ov:*
21: em_pat:*
22: em_ph:*
23: em_pl:*
24: em_ro:*
25: em_sts:*
26: em_sy:*
27: em_un:*
28: em_vl:*
29: gb_ba1:*
30: gb_in1:*
31: gb_in2:*
32: gb_in3:*
33: gb_pl3:*
34: gb_pr4:*
35: em_ba1:*
36: em_ba2:*
37: em_hg1:*
38: em_hg2:*
39: em_hg3:*
40: em_hg4:*
41: em_hg5:*
42: em_hg6:*
43: em_hg7:*

44: em_hg8:*
45: em_hg9:*
46: em_hg10:*
47: em_hum3:*
48: em_hum4:*
49: em_hum5:*
50: em_hum6:*
51: gb_pr5:*
52: gb_pr6:*
53: gb_pr7:*
54: gb_hg1:*
55: gb_hg2:*
56: gb_hg3:*
57: gb_hg4:*
58: gb_hg5:*
59: gb_hg6:*
60: gb_hg7:*
61: gb_hg8:*
62: gb_hg9:*
63: gb_hg10:*
64: gb_hg11:*
65: gb_hg12:*
66: gb_hg13:*
67: gb_hg14:*
68: gb_hg15:*
69: gb_hg16:*
70: gb_hg17:*
71: gb_hg18:*
72: gb_hg19:*
73: gb_hg20:*
74: gb_hg21:*
75: gb_hg22:*
76: gb_hg23:*
77: gb_hg24:*
78: gb_hg25:*
79: gb_hg26:*
80: gb_hg27:*
81: gb_hg28:*
82: gb_hg29:*
83: em_hg0:*
84: gb_hg24:*
85: gb_hg28:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2905	100.0	2905	10	AF084645 Homo sapi
2	2807.6	96.6	4448	52	HSJ9936 Homo sapi
3	2651.2	91.3	2802	10	AF084644 Homo sapi
4	2575.6	88.7	4337	52	HSJ9937 Homo sapi
5	2086.2	71.8	2146	10	AF061056 Homo sapi
6	1398.8	48.2	160185	69	AC069444 Homo sapi
7	1100.4	37.9	1885	3	AF182217 Oryctolagus
8	1006.8	34.7	1709	11	AF031814 Mus muscu
9	1004.2	34.6	1755	11	AF151377 Rattus no
10	991.4	34.1	1601	3	AF188476 Oryctolagus
11	335	11.5	365	77	G36929 SHGC-56597
12	284	9.8	1340	4	AF276753 Gallus ga
13	280.8	9.7	1377	11	MUSVDR
14	278.8	9.6	2043	11	RATDHYRDS
15	275	9.5	2484	4	AF011356 Gallus ga
16	272.8	9.4	1284	10	AF026260 Homo sapi
17	272.8	9.4	1335	53	HSVD3R
18	272.8	9.4	4604	85	HUWVDR
19	267.8	8.9	1724	4	CJ012641 Columba ja
20	258.6	8.9	1638	4	XLRANONST
21	239.8	8.3	2429	4	AB037674 Paratrich

22	236.6	8.1	1503	4	AF164512	AF164512	Danio rer
23	236	8.1	1071	81	E14584	E14584	Rat mRNA
24	232	8.0	1782	4	U91846	Xenopus lae	
25	230.4	7.9	2659	4	AB037673	Paralichth	
26	188.6	6.5	1404	81	E14585	Human mRNA	
27	172.2	5.9	1450	53	HSONHORE	230425 H. sapiens m	
28	170.6	5.9	1450	81	AR009748	Sequence	
29	170.6	5.9	1450	82	I73479	Sequence 1	
30	170.6	5.9	1450	82	I81188	Sequence 1	
31	168.6	5.8	1350	11	AF009327	AF009327 Mus muscu	
32	154.2	5.3	1077	11	AF133094	AF133094 Rattus no	
33	152.6	5.3	1077	11	AF133095	AF133095 Rattus no	
34	117.4	4.0	1332	11	AF009328	AF009328 Mus muscu	
35	110.8	3.8	160185	69	AC069444	AC069444 Homo sapi	
36	100.8	3.5	2797	32	DM036792	DM036792 Drosophila	
37	100.8	3.5	113474	58	AC013932	AC013932 Drosophila	
38	100.8	3.5	116280	54	AC007853	AC007853 Drosophila	
39	100.8	3.5	161891	54	AC008206	AC008206 Drosophila	
40	100.8	3.5	227398	80	AF003750	AF003750 Drosophila	
41	100.4	3.5	203	85	HM04DRPBM	M65208 Human vitam	
42	97.6	3.4	1898	81	E14584	E14584 Sequence 1	
43	97.6	3.4	1979	21	E1456	E1456 cDNA encodi	
44	97.6	3.4	1979	81	AR035536	AR035536 Sequence	
45	97.6	3.4	2010	53	HS07132	HS07132 Human stero	

ALIGNMENTS

RESULT 1

LOCUS AF084645 2905 bp mRNA PRI 20-OCT-1998
 DEFINITION Homo sapiens orphan nuclear receptor (PAR1) mRNA, complete cds.
 ACCESSION AF084645
 VERSION AF084645.1 GI:3769538
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 2905)
 AUTHORS Bertilsson,G., Heidrich,J., Svensson,K., Asman,M., Jendeborg,L.,
 Sydow-Backman,M., Ohlsson,R., Postlund,H., Blomquist,P. and
 Berkenstam,A.
 TITLE Identification of a human nuclear receptor defines a new signaling
 pathway for CYP3A induction
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 95 (21), 12208-12213 (1998)
 MEDLINE 98445350
 REFERENCE 2 (bases 1 to 2905)
 AUTHORS Bertilsson,G., Asman,M., Blomquist,P. and Berkenstam,A.
 TITLE Direct Submission
 JOURNAL Submitted (19-AUG-1998) Cell and Molecular Biology, Medical Nobel
 Institute, Karolinska Institute, Doktorsringen2, Stockholm 17177,
 Sweden

FEATURES

gene
 CDS

1. .2905
 /location="Qualifiers
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /tissue_type="Liver"
 1. .2905
 /gene="PAR1"
 280. 1584
 /gene="PAR1"
 /codon_start=1
 /product="orphan nuclear receptor"
 /protein_id="AAC6458.1"
 /db_xref="GI:3769539"
 /translation="MEYRPSKSNHADVHCEDESYGKPSVNADEEGVQICRRC
 GKATGYHFNMTGCGGKGFRRAMKRNARLRCFRRGACETIRKTRROOACRLRC
 LSEGRKKKINSDAVERRRRLIRKRSERCTGPVOCITLROPMITLMDADOK
 TFDFTSHKFNRLPGVLSGCELPESLQASREKAKNSQVRDLCLSLVSLQKRE
 DGSVWNNKPPADSGKRTFSLPDMADSTYMGIIISPAKVISYFEDLPIEIOISLL
 KGAAFEICOLRFNIVFNAETGTWECGRSLCYLIDTAGFQDLLEPMKPKHYMKIKIO

LMHEEYVIMQAIISLSPDRPGVLOHRYVDLOJEOFAITLKSYIECNRPQPAHRLFLK
 IMATLELRSLNHOHRLRIDIDHFAPIMLQELFGITGS"
 BASE COUNT 765 a 727 c 778 g 635 t
 ORIGIN

Query Match 100.0%; Score 2905; DB 10; Length 2905;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 2905; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	cctctgaaggttcctgaatcgatagtagtaattcgtggacgggaaggaaggaagcactgcct	60
DB	1	CCTCTGAAGGTTCTGAATCGATAGTAGTAATTCGTGGACGGGAAGGAAACACTGCTT	60
QY	61	tacttcagtggaatcctgcgcctcagccttggaagcaaggtttcaacagtgaagaaagca	120
DB	61	TACTTCAGTGGAAATCTGCGCTCAGCTCGAAGCAAGTGTCACTGAAAAGCA	120
QY	121	ggaataagc'ataactcctgcctgaacaaagcagcgctccttgtaaaagctactcct	180
DB	121	GGAATAAGCTAATACTCTCTCTCGAACAAGCAGCGGCTCTGTAAGCTACTCT	180
QY	181	tgatcgatcctttgacacggatgttcaaaagtggaccccaagggaagatcgtggacaag	240
DB	181	TGATCGATCTTGGACCGGATTTGTTCAAAAGTGGACCCCAAGGGAAGTGGAGCAAG	240
QY	241	aacttaccacaaagagtgccaaagagcccaagaaacccctggaggtgtagacccaagaa	300
DB	241	AACCTTACCACAAAGAGTGCCAAAGAGCCCAAGAAACCTGGAGGTAGACCCAAAGAA	300
QY	301	agctgaaacatcgtgactttgtacactgtgagacacagaagtctgtccctggaaagccc	360
DB	301	AGCTGAAACATCGTACTTGTACACTGTGAGACACAGAGTGTCTCGAAGAGCCC	360
QY	361	agtgtaacgcagataggaagatcggaaggtccccaactcgcgtgtatgtggygacaag	420
DB	361	AGTGTAACGAGATGAGGAAGTCCGAGGTCCCAACTCGCTGTATGTGGGACAG	420
QY	421	ggcactggctatcactcaatgtagtagatgtagaagatgaaagggcttttcaggagg	480
DB	421	GGCACCTGGCTATCACTTCAATGTGATGATGATGAAGATGCAAGGCTTTTTCAGGAG	480
QY	481	ggcatgaaacgcagacgcggcgctgaggtgcccttcggaaagggcgctcgagataacc	540
DB	481	GGCATGAACGCAAGCCGCGCTGAGGTGCCCTTCGGAAGGCGCTGCGAGATCAC	540
QY	541	cggaaagaccggcggaagatgccaaggtcgccgcctggcggaagtgcttggaagggcagt	600
DB	541	CGGAAGACCCGGGACAGTGCAGGCGCTGCCCTGCGCAAGTGCCTGGAGACGGCAGT	600
QY	601	aagaagagatgatactgtccagacggcgctgtagaaggaagggcgagccttgatacaagcg	660
DB	601	AAGAAGAGATGATCTATGTCCGACAGAGCGCTGAGAGGAGAGGCGGCTTGATCAAGGG	660
QY	661	aagaagaagtgaacgcagcaggaagcactgggaagtgcaggggtcagagaaggaagcaag	720
DB	661	AAGAAAGTGAACGACAGAGGACTCAGCCACTGGAGTGCAGAGGCTGCAAGAGAGAG	720
QY	721	cggatgatatacaggaagatgtgagcgtcagatgaanaaccttgaacatacctctcc	780
DB	721	CGGATGATGATCGGAGAGTGTAGAGCTCAGATGAANAACCTTGACACTTCTTCC	780
QY	781	cattcaagaattccggcgtcgcaaggggtcttaagcagtgtcgtgagttgcagaagttc	840
DB	781	CATTCAAGAATTCGGGCTGCCAGGGGCTTTAGAGAGTGTGAGAGTGCAGAGTCT	840
QY	841	ctgcagggcccatcgaggggaagaagctgcgaagtggagccaggtccogaaagatctgtgc	900
DB	841	CTGCAAGGCCCATCGAGGGGAAGAGCTGCCAAGTGAAGCCAGAGTCCGGAAATCTGTGC	900
QY	901	tctttgaaggtctctctgcaagctcgggggggaagatggagatgctggaactacaacc	960
DB	901	TCCTTGAAGGTCTCTCTGCAAGCTCGGGGGGAGAGATGGAGATGCTGGAACCTACAAACC	960

```

Oy 961 ccaagccagctgagcgggaaagagatctctccctgctgccccacatggtcagatgca 1020
Db 961 ccaagccagctgagcgggaaagagatctctccctgctgccccacatggtcagatgca 1020
Oy 1021 acctacatgttcaaaaggaatcactgcttgcacaaatgcaatctctacttcaaggacttg 1080
Db 1021 acctacatgttcaaaaggaatcactgcttgcacaaatgcaatctctacttcaaggacttg 1080
Oy 1081 cccatcgaagagacagatctcctgctgaaaggaggcgcgttccgagctgtgtcaactgaga 1140
Db 1081 cccatcgaagagacagatctcctgctgaaaggaggcgcgttccgagctgtgtcaactgaga 1140
Oy 1141 ttcacacacagtggttcaacgcygagacatggaactgtgagctgtgagcctgctcactgac 1200
Db 1141 ttcacacacagtggttcaacgcygagacatggaactgtgagctgtgagcctgctcactgac 1200
Oy 1201 ttggaagacacacgagctgagcttccagcaacttctactgagagccatgctgaatccac 1260
Db 1201 ttggaagacacacgagctgagcttccagcaacttctactgagagccatgctgaatccac 1260
Oy 1261 tacatgtgaaagacgctgagctgcaatgagagagtagtgctgtagtcagggcactcctc 1320
Db 1261 tacatgtgaaagacgctgagctgcaatgagagagtagtgctgtagtcagggcactcctc 1320
Oy 1321 ctctctcccccagacccgcccaggtgtgtgcaagacccgctgtgtgagacacgctgcaaggag 1380
Db 1321 ctctctcccccagacccgcccaggtgtgtgcaagacccgctgtgtgagacacgctgcaaggag 1380
Oy 1381 caatcgcacatctactgcaagcttcaactgaaatgcaatcggccacgctgctcaatgag 1440
Db 1381 caatcgcacatctactgcaagcttcaactgaaatgcaatcggccacgctgctcaatgag 1440
Oy 1441 tctctgttcctgaaagatcagctgtagtcaacggagctccgagagataaatgctcagac 1500
Db 1441 tctctgttcctgaaagatcagctgtagtcaacggagctccgagagataaatgctcagac 1500
Oy 1501 acccagcggctgtgtgcaatccagagacatacacccttgcctgtaagccctcactgcaaggag 1560
Db 1501 acccagcggctgtgtgcaatccagagacatacacccttgcctgtaagccctcactgcaaggag 1560
Oy 1561 ttgttcgagcatcaagagtagctgagcggctgctcctgtgtgcaacctccgagagagagc 1620
Db 1561 ttgttcgagcatcaagagtagctgagcggctgctcctgtgtgcaacctccgagagagagc 1620
Oy 1621 agaccacagcgcctctgagcgcgcacctccgggccaagaagatgagacagcgcgaagagc 1680
Db 1621 agaccacagcgcctctgagcgcgcacctccgggccaagaagatgagacagcgcgaagagc 1680
Oy 1681 cgaacaatgcccctgtgctgtctcctcctgaggaattcctgctatgacagcctgctagacat 1740
Db 1681 cgaacaatgcccctgtgctgtctcctcctgaggaattcctgctatgacagcctgctagacat 1740
Oy 1741 tccctaaggaagagacatggtgtgccccccacccacagctcaagctgtgaagagtgaaagccac 1800
Db 1741 tccctaaggaagagacatggtgtgccccccacccacagctcaagctgtgaagagtgaaagccac 1800
Oy 1801 agactcttaagctgagagagtgacatgtagtcaagagacatgagagagcaaggttg 1860
Db 1801 agactcttaagctgagagagtgacatgtagtcaagagacatgagagagcaaggttg 1860
Oy 1861 ccccttccctttaaagagccctgtgtgtgtgctgaggaagaaatccctcagaccccaataagtg 1920
Db 1861 ccccttccctttaaagagccctgtgtgtgtgtgctgaggaagaaatccctcagaccccaataagtg 1920
Oy 1921 tcaaggtgtgagagagacccaagagcaagagatagagcactgtgggtctctatgccccacata 1980
Db 1921 tcaaggtgtgagagagacccaagagcaagagatagagcactgtgggtctctatgccccacata 1980
Oy 1981 ccacaggtgtgtgctgctgagcttctcaatgctactgctcaataagcctgctccacac 2040
Db 1981 ccacaggtgtgtgctgctgagcttctcaatgctactgctcaataagcctgctccacac 2040

```

```

Oy 2041 ttcccaactgcttcccccctcctctccagagctgctgtgtgggtcgaagcctgactcatcg 2100
Db 2041 ttcccaactgcttcccccctcctctccagagctgctgtgtgggtcgaagcctgactcatcg 2100
Oy 2101 gcaaggtgcatgagatctctgtggagctcctctagaagatgagaagcgaagagcctgacac 2160
Db 2101 gcaaggtgcatgagatctctgtggagctcctctagaagatgagaagcgaagagcctgacac 2160
Oy 2161 caaatgtcgaagaaagcttggcctgagcctcattccgggacacatctctgtctctgatacc 2220
Db 2161 caaatgtcgaagaaagcttggcctgagcctcattccgggacacatctctgtctctgatacc 2220
Oy 2221 attggaacacatattaaagacatgataaagtagtcctgtgtggtgatacagacatga 2280
Db 2221 attggaacacatattaaagacatgataaagtagtcctgtgtggtgatacagacatga 2280
Oy 2281 ctgagatatagatactcgtgagctcagagagttatagtttaaaaaaaacaagaacaacaa 2340
Db 2281 ctgagatatagatactcgtgagctcagagagttatagtttaaaaaaaacaagaacaacaa 2340
Oy 2341 caatttgatcaaaagagagaaatgaatgaatgaatgaatgaatgaatgaatgaatgaatga 2400
Db 2341 caatttgatcaaaagagagaaatgaatgaatgaatgaatgaatgaatgaatgaatgaatga 2400
Oy 2401 gtgagatgctgagctgtgagtgagcagcctggtgtacccaaagtgaaggttcccgagagacatg 2460
Db 2401 gtgagatgctgagctgtgagtgagcagcctggtgtacccaaagtgaaggttcccgagagacatg 2460
Oy 2461 agtctgtgaagagcaagggcacaacacgtgagctgtgagctgtgtgtgtgtgtgtgtgtgt 2520
Db 2461 agtctgtgaagagcaagggcacaacacgtgagctgtgagctgtgtgtgtgtgtgtgtgtgt 2520
Oy 2521 gttaggtctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2580
Db 2521 gttaggtctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2580
Oy 2581 tgcctgtgcaagagctgagctgagcagcagctgagcagcagcagcagcagcagcagcagcag 2640
Db 2581 tgcctgtgcaagagctgagctgagcagcagcagcagcagcagcagcagcagcagcagcagcag 2640
Oy 2641 tgcacattatctctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2700
Db 2641 tgcacattatctctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2700
Oy 2701 cctgtgttaagcagcattgtgagtaaaatttttgcacatttcaaaatttacttct 2760
Db 2701 cctgtgttaagcagcattgtgagtaaaatttttgcacatttcaaaatttacttct 2760
Oy 2761 ataagcattccacacacacacacacacacacacacacacacacacacacacacacacacac 2820
Db 2761 ataagcattccacacacacacacacacacacacacacacacacacacacacacacacacac 2820
Oy 2821 atcaagcgaagaaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatga 2880
Db 2821 atcaagcgaagaaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatga 2880
Oy 2881 aaaaaaaaaaaaaaaaaaaaaaaaaa 2905
Db 2881 aaaaaaaaaaaaaaaaaaaaaaaaaa 2905

```

RESULT 2
 HSAJ9936 4448 bp mRNA PRI 08-SEP-1999
 LOCUS
 DEFINITION Homo sapiens mRNA for nuclear hormone receptor PRL.
 ACCESSION AJ009936.1 GI:5852062
 VERSION
 KEYWORDS nuclear hormone receptor; orphan nuclear receptor; PRL gene.
 SOURCE human
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 4448)

AUTHORS Bertilsson, G., Asman, M., Blomquist, P. and Berkenström, A.
 TITLE Direct Submission
 JOURNAL Submitted (19-AUG-1998) Cell and Molecular Biology, Medical Nobel
 Institute, Karolinska Institute, Doktorsringen 2, Stockholm 17177,
 Sweden

FEATURES

source

Location/Qualifiers

1. 2802
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /tissue_type="Liver"
 1. 2802
 /gene="PAR2"
 60. 1481
 /gene="PAR2"
 /codon_start=1
 /product="Orphan nuclear receptor"
 /protein_id="AAC64557.1"
 /db_xref="GI:3769537"

CDS

/translation="MTVTRTHHFKGSLRAPALPLHSAALAEIASHNRPANLEVR
 KESNHADEVHEDESYVPRKPSVMADEVGPOICRVCGDKATGAPHWYTCGCG
 FPRAMKRNARLRCPFRKRGACETIRKTRROQARLRKLCESGKMKIMSDAEVER
 RALIKRKSERTGIDPLGVGLIEBQRNMLRELMADQKTFDTTFSHKMFRLGVLS
 SGCELPESLQAPSRBAKWSQVRDLCSLVSQLRGDSVWNYRPPADSGKEFT
 SLPHMADSTYMFKGLISPAKVISYFDLPLEIQLSLKAAEELCOLPNTYFNE
 TGTWECGRSLSYCLEDTAGFOQLLEPLKFKHYMLKILQLEEVYVLMQISLESPR
 PGVLOHRYVDLOBOFAITTKSYIECNRPQAHFPLFKIMMLTELRSINAOHTQRL
 LRIDIHFAPIPMDELFTIGTS"

BASE COUNT 723 a 715 c 755 g 609 t
 ORIGIN

Query Match 91.3% Score 2651.2; DB 10; Length 2802;
 Best Local Similarity 99.7%; Pred. No. 0;
 Matches 2656; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

242 acttaccacaaagcagtcacaaaggcccaagaagcaaacctggaggctgaaccacaaagaa 301
 Db 139 AGTTGGCTTCAAAACATCCAAAGAGGCCCAAGCAAACTGGAGGTGAGACCCAAAGAA 198
 QY 302 gctggaacatgctgactcttgtaactgtgaggaacacagatcttccctggaagccca 361
 Db 199 GCTGGAACCATGCTACTTGTACACTGTGAGGACACAGAGCTCTTCTTGAAAGCCCA 258
 QY 362 gtgtcaacgcagatgaggaagtcgagagtcaccaatctgcgcgtgtatgtgaggacaag 421
 Db 259 GTGTCAAGCAGATGAGGAGTCGAGGATCCCAAAATGTCGCGTATGTCGGGACAAAG 318
 QY 422 ccaatgctatcaactcaatgtcatgtgaagatgtgaagggctttttcagaaggg 481
 Db 319 CCACGTGCTATCACTTCAATGATGATGATGAGATGCAAGGAGGCTTTTTCAGAGAGG 378
 QY 482 ccaatgaacgcagacgc 541
 Db 379 CCATTAACCGCAACCCCGGCTGAGAGTCCCTTCGGAAAGGCGCCCTGCGAGATCCAC 438
 QY 542 ggaagaccgc 601
 Db 439 GGAAGACCCGGGACAGTGCAGAGGCGCTCGCTGCGAATGCTGAGAGCGCGCATGA 498
 QY 602 agaagagatgatactgtcgaagagcgcgtggaagagagcgagccttgatcaagcgga 661
 Db 499 AGAAGGAGATGATCATGTGCCACGAGGCGCTGAGAGAGCGGCGCTTTCAGAGCGGA 558
 QY 662 agaaagtgagacgc 721
 Db 559 AGAAAGTGAACGACAGAGGACTACGCACTGGAGGTGTCAGAGGCTGACAGAGGAGAGC 618
 QY 722 ggaatgatcatcaggaagctgtgagcgcctcagatgataaacccttgacactcctccccc 781
 Db 619 GGATGATGATCAGGAGGAGTGTGAGCGCTCAGATGAAGAACTTTGACACTTACTTCTCCC 678
 QY 782 attcaagaattccgc 841

Db 679 ATTCAAGAAATTTCCGGCTGCCAGGSGCTGTAGCACTGCTCCGAGTTGCCAGACTTC 738
 QY 842 tgcagcccccctcgaagggaagagctgccaagtggagccaggtccggaagatctgtgtc 901
 Db 739 TGCAGGCCCATCGAGGGAAGAGCTCCAAAGTGGAGCCAGGTCCGGAAGATCTGTGCT 798
 QY 902 ctgtgaaggtctctctcgaactcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 961
 Db 799 CTTTGAAGGCTCTCTCTCAGACTCGGGGGAGAGTGGCAATGCTGTGAACTTAAACCCC 858
 QY 962 cagccgcagcagtgccggaagagatctctccctgctgcgcgcgcgcgcgcgcgcgcgcgc 1021
 Db 859 CAGCCGACAGTGGCGGGAAGAGATCTTCTCTCTGCTGCCCAATAGGCTGACATGTCAA 918
 QY 1022 cctacatgttcaaaaggcatcaatcagctttgccaagtcatctcctaactcagggacttgc 1081
 Db 919 CCAACAGTTTCAAAAGCATCATCAGCTTTCCAAACTCATCTCTCACTCAGGAGACTTGC 978
 QY 1082 ccatcgagagacatctcctctcgtgaaggggcgcgccttcgaagctgttcaactgaat 1141
 Db 979 CCATCGAGGACCAAGATCTCTCTGTAAGGGGGCCCTTCGAGCTGTCAACTGAGAT 1038
 QY 1142 tcaacacagtgctcaacgc 1201
 Db 1039 TCAACACAGTTCACACGCGGAGACTGGAACCTGGGAGTGTGGCCGCTGTCTACTGCT 1098
 QY 1202 tggagaacacacagtgagtgcttcccaacattctacttggagcccatgtgaattccact 1261
 Db 1099 TGGAAACACTGACGAGGTGCTTCCAGCAACTTCTACTGGAGCCCATGCTGAATTCACAT 1158
 QY 1262 acatgctgaaagagctgcaagctgcatgagaagagatgtgtatgagagccatctccc 1321
 Db 1159 ACATGCTGAAGAGAGCTGACGCTCATAGAGAGAGATGTGCTGATGACGCGCATCTCCC 1218
 QY 1322 tcttctcccaacgc 1381
 Db 1219 TCTTCTCCCGACACCCCGCAGGTGTCTGACAGACCGGTGTGAGCACTGACGAGAAC 1278
 QY 1382 aattgcacatctctgaagctcctacattgaatgaatcggcccaagcctgtcataggt 1441
 Db 1279 AATTGCGCATTACTCTGAAGTCTTACATTAATGCAATGGCCCAAGCCTGCTCATAGGT 1338
 QY 1442 tctgttccctgaagatcatgagctatgctacacagagctccgcagatcaaatgtcgaca 1501
 Db 1339 TCTTGTCTCTGAAGATCATGCTATCTACCGAGTCCGACATCAATGCTCAGCA 1398
 QY 1502 cccaagcgtctgctgcacatccaggaacatacaccccttctacgcgcctcatatgcagagt 1561
 Db 1399 CCCAGCGGCTGCGCATCCAGAGACATACACCCTTGTCTACGCGCCCTCATGACGAGAT 1458
 QY 1562 tgttggcatcaaggtatcgaagcgtgccttgggtgacacctcgcagagcgagcca 1621
 Db 1459 TGTTCGGCATCAAGGTATCTGAGCGGTGCGCTTGGGTGACACCTCCGAGAGGAGCA 1518
 QY 1622 gaccagagccctctgagc 1681
 Db 1519 GACCCAGAGCCCTCTGAGCGCGCCACTCCCGGSCCAAGACATGACATGCCAAGAGACC 1578
 QY 1682 gacaatgcctctgctgcctctcctcgaaggaattcctctgtaagaactgtgtacatt 1741
 Db 1579 GACAATGCCCTGCTGCGCTCTCCCTAGGAAATCTGCTATGACAGCTGGCTATGACT 1638
 QY 1742 cctcaggaagagatgagtggtgcccccccaaccagttcagctgtgtaggaagtgaagcaca 1801
 Db 1639 CCTCAGGAAGGACATGGGTGCCCCACCCACAGTTCACTGTGTGAGAGTGAACCCACA 1698
 QY 1802 gactcttaagtagagtgactgacactgttagtgtaaggaacataaggaagcaagtggtgc 1861
 Db 1699 GACTCTTAGTGGAGAGTCACTGACTGTAGAGTACAGACATCAGAGAGCAAGCTTGC 1758
 QY 1862 ccttctctttaaagggccctgtgtctggtgggagaaatccctcagatcccaataaggtgt 1921
 Db 1759 CCTTCTCTTTTAAAGGCCCTGTGTGTGGGAGAAATCCCTCAGATCCCACTAAAGTGT 1818

Oy	1922	caaggctgaggagagcaaaagcgaagagatgagccatctgaggctgtctgtccacataac	1981
Db	1819	caagctgtgcaaaagacacacacgacacaaagatgagccatctgaggctgtgtgccacataac	1878
Oy	1982	ccagctgtgtctgcgtctccctgagctctttctcatctgtacaccttaatagctcgtctccacat	2041
Db	1879	ccagctgtgtgttcgcttccctgagctctttctcatctgtgtacaccttaatagctcgtctccacat	1938
Oy	2042	tcccaactcgctcccccctccctctctccgagctgcttgtgtggctcaaggctctgtactcaacg	2101
Db	1939	ttcccaactcggttccccctccctctctccgagctgcttgtgtggctcaaggctctgtactcaacg	1998
Oy	2102	caaggctgcaatgatactctgtggagtgctctctcagaagagatgaaagccagaagagccgcacac	2161
Db	1999	caaggctgcaatgagatctgtggagtgctctctcagaagagatgaaagccagaagagccgcacac	2058
Oy	2162	aaatgtcagaagagcttggacatgacccatctccggccacacatctctgtgtctctgcatcca	2221
Db	2059	aaatgtcagaagagcttggacatgacccatctccggccacacatctctgtgtctctgcatcca	2118
Oy	2222	tttggacaacatattatgaagcactgttaataagtgctgtctgtgtgggtatacagcaatgac	2281
Db	2119	tttggacaacatattatgaagcactgttaataagtgctgtgtgtgggtatacagcaatgac	2178
Oy	2282	tcaagatacagaacctgtgacgacacagagctttatatagcttaaaaaaacaacagaaacacacac	2341
Db	2179	tcaagatanaagatcctgagagctcacagagattatagttaaaaaaacaaacagaaacacacac	2238
Oy	2342	aacttggatcacaagagagaaatgataagtgacaaagaacagacagaagaaattccctgtg	2401
Db	2239	aatttggatcamaaaagagaaatgataagtgacaaagaacagacagaagaaattccctgtg	2298
Oy	2402	tggatgtcgagagctgtgtagtgagcagagcactgtgtacccaagtgaagttcccgagacatga	2461
Db	2299	tggatgtcgagagctgtgtagtgagcagagcactgtgtacccaagtgaagttcccgagacatga	2358
Oy	2462	gtctgtgagagcaagagagcaaaaactgcagctgtgagctgctgtgtgtgatttggtgtag	2521
Db	2359	gtctgtgagagcaagagagcaaaaactgcagctgtgagctgctgtgtgtgatttggtgtag	2418
Oy	2522	taagttctgttgcactgtatgtgggcccgtgggttgtctccctgggctgtgaatgtcggat	2581
Db	2419	taagttctgttgcactgtatgtgggcccgtgggttgtctccctgggctgtgaatgtcggat	2478
Oy	2582	gctctgtgacaagagctacgctgtgacaactcagattaaacacccggagagaaacatttcat	2641
Db	2479	gctctgtgacaagagctacgctgtgacaactcagattaaacacccggagagaaacatttcat	2538
Oy	2642	gcacattatattctgtgtacacatctatctcacaagctcaaaaggtacgaagaagctgtgc	2701
Db	2539	gcacattatattctgtgtacacatctatctcacaagctcaaaaggtacgaagaagctgtgc	2588
Oy	2702	ctgttctatgtccacactgtgtagttaaaaaatttttttgcatcttcacaaatctacttata	2761
Db	2599	ctgttctatgtccacactgtgtagttaaaaaatttttttgcatcttcacaaatctacttata	2658
Oy	2762	taagagcaatccacaccccaagaagctagttttgtggaaatgaaagccgcgggttaaatgtcaaa	2821
Db	2659	taagagcaatccacaccccaagaagctagttttgtggaaatgaaagccgcgggttaaatgtcaaa	2718
Oy	2822	tcaagagcaaaaggaatbaaaataatgtacttttgcttaaaaaaataaaaaaataaaaaa	2881
Db	2719	tcaagagcaaaaggaatbaaaataatgtacttttgcttaaaaaaataaaaaaataaaaaa	2778
Oy	2882	aaaaaaaaaaaaaaaaaaaaaaaaa 2905	
Db	2779	aaaaaaaaaaaaaaaaaaaaaaaaa 2802	

DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	REFERENCE	AUTHORS	TITLE	JOURNAL	FEATURES	SOURCE	
Homo sapiens mRNA for nuclear hormone receptor PRR2.	AF009937	1	GI:5852066	nuclear hormone receptor; orphan nuclear receptor; PRR2 gene; human.	Homo sapiens	Ekayacya; Metzosa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	1 (pages 1 to 4337)	Heard,D.J., Holloway,J., Hansen,C., Tommerup,N., Aagaard,L. and Vissing,H.	Identification of a novel protein isoform of the human nuclear hormone receptor PXR/SXR and localization to chromosome 3q12.1	Eur. J. Hum. Genet.	In press	2 (pages 1 to 4337)	Heard,D.J.	Direct Submission	Submitted (31-JUL-1998)	Heard D.J., Novo Nordisk A/S, Molecular Genetics, Novo Alle, DK-2880, Bagsvaerd, DENMARK
							1.4337	Location/Qualifiers								

gene
CDS

```
/gene="PRK2"  
/note="isoform with longer A/B domain, translation  
initiates at GUG-1771 5 fold less efficiently than at  
CUG-1840"
```

/codon_start=1
 /transl_except=(pos:1771..1773,aa:Met)
 /product="nuclear hormone receptor PR2-C"
 /protein_id="CAB55493.1"
 /db_xref="gi:5852068"
 /translation="MDPRGEVGAKLPPNSRPEANLEVRPKSWMHTDVFHCEDTET
 STVPEPSVNADEEVGGPOICVCGDKAGVYHFWTCGCGKGFRRRAKRNARLRCPPE
 RKGAQEIETRRKDOQARLKRKCESGKFNINSDEVAERBRLAIKKRKEGTGP
 LGVQDIEQPMRIELMDAQMKETDFTFSFKFNRSVLQJREGESVWYKSPADGTO
 GKEISLIPHNADMTYFEGKISPAFYISYRDLPTIDQSLIKGAFAECLQIRPNP
 YENATGMEGRGSLCYCEDTAGCGOOLLEPMLKFMHMKKLLQHEEYLYLMAIS
 PSPDPGLOHRVYDQJQEQFAITKSTIECNRFPQARHFLKLMALTELSINMO
 TQRLRLKRDTHPRFATPLMQELFGITGS"
 :1840..3033

```

/gene="PRR2"
/function="transcription factor"
/note="full length receptor"
/codon_start=1
/product="nuclear hormone receptor PRR2-A"
/protein_id="CAB5492.1"
/db_xref="GI:5852067"
/translation="MEVPRKESNMHTDFVHCEDTESVPGKPSVNADEVGGPQICRVC
GDKATGYFNTVCCECKGFPRPAAKRNARLRCPFGACITRTTRROCAQRLRC
TESGKNRNMISDAVEERBALIKRKSESGTQPLGVGLTSPDPMIIRLMDAOKK
TEDTFSEFKNRSILORGDSGVNMYKSPADSGKKEISLPHMDMSYMKGTIT
SEAVISFRLDPIEDQISLKGAFAELCQNRFTYVAEKGTECGRLSTCEDTNG
GFQOILLEPRALFKHYMKKKLDLHEEYVLMQALISESDRGVILQHNVODLQCFM
TLKSTIECNRRQPAHRLFKLIMATLTSLRSINQHTORLLIDDIHPFAPLMQELE
GITGS"
2005..3033

```

RESULT	4
HSNJ9937	
LOCUS	
HSNJ9937	4337 bp
	mRNA
	PRI
	09-SEP-1999
<pre> /product="nuclear hormone receptor PRR2-B" /protein_id="CAB55494.1" /db_xref="GI:5852069" /translation="MTCEGCKGFFBRAKRNAAALRCPPRKACETITKTRHQCACNL </pre>	

[illegible]

FEATURES	source
AUTHORS	Lehmann, J.M., McKee, D.D., Watson, M.A., Willson, T.M., Moore, J.T. and Kiewer, S.A.
TITLE	The human orphan nuclear receptor PXR is activated by compounds that regulate CYP3A4 gene expression and cause drug interactions
JOURNAL	J. Clin. Invest. 102 (5), 1016-1023 (1998)
MEDLINE	98395173
REFERENCE	2 (bases 1 to 2146)
AUTHORS	Lehmann, J.M., McKee, D.D., Watson, M.A., Willson, T.M., Moore, J.T. and Kiewer, S.A.
TITLE	Direct Submission
JOURNAL	Submitted (21-APR-1998) Molecular Endocrinology, Glaxo Wellcome, 5 Moore Drive, RTP, NC 27709, USA
FEATURES	Location/Qualifiers
source	1..2146
CDS	/organism="Homo sapiens" /db_xref="taxon:9606" /tissue_type="Liver" 304..1608 /codon_start=1 /product="Orphan nuclear receptor PXR" /protein_id="AAD05436.1" /db_xref="gi:3511138" /translation="MEYRPKESRNHADFDVHCEDESYPGRFSVNADEVGPOICRVCGDKATGHFNVTCEGCGEFGFRAMKRNARLRCPFRGACITRTKTRCOACARLRCLESMSKEMIMSDIAVEBERRALIRKRSERGTGLOVGOQTEEDQMLRELDQOMHITFDTFSHFNKRLRPGVLSGCELPESLOAPSRREAAWSOVRDLCSLXVLSLQDRGDSVYNNKPPADSGSGKEIFSLPPHADMSTYFMGILSFPAVISYFPDIPEDDISILLKGAELCOLLENTVFNAETGTWEGCSLSTLEPTAGGSFOILLERPHKRYHLKKLQ LHEEYVIMQAI SLFSPDRPVQHRVYDQLOEOPATLLKSYIECNRPQPAHRLFLTKLIMATLETRSLNMAHRTORLRIODIHFPAPFPMQELFGITGS"
BASE COUNT	509 a 593 c 605 g 439 t
ORIGIN	
Query Match	71.8%; Score 2086.2; DB 10; Length 2146;
Best Local Similarity	99.4%; Pred. No. 0;
Matches 2094; Conservative	0; Mismatches 13; Indels 0; Gaps 0;
16	gaatcgatagtagtaattcgctggagcgggaagaagaagcactgcttacttactagtaggaat
40	GAAATCATTAACCTTAGATGAGTAGGAGGGAAGGAAGCACTGCTTACTTACTGATGGGAAT
76	ctcggccctcagcctcgcgaagcgaagtgttcacaaagtgaaataaagcaagagaataagctaata
100	CTCGGCTCAGCCTGCGCAAGCCAAAGTGTCAACAGTAGAAGAAAGCAAGAGAAATTAAGCTATA
136	ctctcgtctcgaacaaagcagcggctccttctgttaagctactccttgaatcgatccttgc
160	CTCCGTCTCTGAACAAGGACGCGGCTCCTTGTAAGCTACTTCTTATTCATCTCTTTC
196	accggatgttccaaagtgagaccccaaggggagagaagtctggagcaagaacttaccacaagc
220	ACCGGATGTGTCAAAGTGAGACCCCAAGGGGAATACTCGAGCAAAAGAACTTACCCACCAAGC
256	agtcacaagaagccccaagaagaacaacctggaagtggagaccacaagaagaagcttgaatacagt
280	AGTCCAAGAGGCCCAAGAAAGCAAACTTGAGAGGTGAGACCCCAAGAAAGTGGAACTGAACTA
316	gaattttagacactgttgaggacacagagtcgtcttcctggaaagcccaagtgttcaacgcagat
340	GACTTTGTACACTGTGAGGACACAGAGTCTGTTCTCTGGAAGCCCAAGTGTCAAGCGAGAT
376	gaggaaatcgggaaggtccccaacatctgcgtgtatgttgagggaacaaagggcaccgtatcac
400	GAGGAATCGGAGGTCCCCAAATCTGCCGTGTATGTGGGACCAAGGCGACCTGGCTATAC
436	ttcaatgtcattgacatgttgagaatgcaaggctttttccaggaggccatgaaacgcgaac
460	TTCAATGTCAATGACATGTGAGATGCAAGGCGCTTTTTCAGGAGGCGCATGCAAGCGCAAC
496	gcccggctggagggtccccctcccgaaaggggcgctgcgagagatacccccggagaacccggcgga
520	GCCCGGTGAGGTGCCCTTCCGGAAGGCGCTTCGAGATGATACCCGGAAGACCCGGCGGA

Direct Submission
Submitted (30-MAY-2000) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Aug 7, 2000 this sequence version replaced gi:8659996.

44-38861-1000

NOTE: Estimated insert size may differ from sequence length (see <http://www.hgsc.bcm.tmc.edu/docs/genbankdraft.data.html>).
NOTE: This is a "working draft" sequence. It currently consists of 34 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

*	1	19705:	contig of 19705 bp in length
*	19706	19805:	gap of unknown length
*	18086	30128:	contig of 10324 bp in length
*	30130	30229:	gap of unknown length
*	30230	40401:	contig of 10172 bp in length
*	40402	40501:	gap of unknown length
*	40502	49848:	contig of 9947 bp in length
*	49849	49948:	gap of unknown length
*	49949	59003:	contig of 9955 bp in length
*	59004	59103:	gap of unknown length
*	59104	64582:	contig of 5479 bp in length
*	64583	64682:	gap of unknown length
*	64683	70787:	contig of 6105 bp in length
*	70788	70887:	gap of unknown length
*	70888	77647:	contig of 6760 bp in length
*	77648	77747:	gap of unknown length
*	77748	83658:	contig of 3591 bp in length
*	83659	83738:	gap of unknown length
*	83739	88544:	contig of 4806 bp in length
*	88545	88644:	gap of unknown length
*	88645	91971:	contig of 3327 bp in length
*	91972	92071:	gap of unknown length
*	92072	98127:	contig of 6056 bp in length
*	98128	98227:	gap of unknown length
*	98228	104128:	contig of 5901 bp in length
*	104129	104228:	gap of unknown length
*	104229	107732:	contig of 3504 bp in length
*	107733	107832:	gap of unknown length
*	107833	112814:	contig of 4982 bp in length
*	112815	112914:	gap of unknown length
*	112915	118277:	contig of 5363 bp in length
*	118278	118377:	gap of unknown length
*	118379	122544:	contig of 4167 bp in length
*	122545	122644:	gap of unknown length
*	122645	125954:	contig of 3310 bp in length
*	125955	126054:	gap of unknown length
*	126055	129656:	contig of 3572 bp in length
*	129657	129726:	gap of unknown length

FEATURES	SOURCE
*	129737
*	132738: contig of 3012 bp in length
*	132739
*	132838: gap of unknown length
*	132839
*	135617: contig of 2779 bp in length
*	135618
*	135717: gap of unknown length
*	135718
*	138538: contig of 2821 bp in length
*	138639
*	138638: gap of unknown length
*	140793: contig of 2155 bp in length
*	140794
*	140893: gap of unknown length
*	140894
*	142848: contig of 1951 bp in length
*	142849
*	142944: gap of unknown length
*	142945
*	145443: contig of 2499 bp in length
*	145444
*	145543: gap of unknown length
*	145544
*	147448: contig of 1905 bp in length
*	147449
*	147548: gap of unknown length
*	147549
*	147599
*	150063: contig of 2515 bp in length
*	150064
*	150163: gap of unknown length
*	150164
*	151748: contig of 1585 bp in length
*	151749
*	151848: gap of unknown length
*	151849
*	153736: contig of 1888 bp in length
*	153737
*	153836: gap of unknown length
*	153837
*	154980: contig of 1154 bp in length
*	154981
*	155080: gap of unknown length
*	155081
*	156398: contig of 1308 bp in length
*	156399
*	156498: gap of unknown length
*	156499
*	157640: contig of 1142 bp in length
*	157641
*	159061: gap of unknown length
*	159062
*	159161: contig of 1331 bp in length
*	159162
*	160185: gap of unknown length
*	160185: contig of 1024 bp in length
Location/Qualifiers	
1..160185	

BASE COUNT	45836	a	32066	c	31244	g	47693	t	3346	others
ORIGIN	.									

Query Match	48.28;	Score 1398.8;	DB 69;	Length 160185;
Best Local Similarity	98.48;	Pred. No. 1.3e-277;		
Matches 1423; Conservative	0;	Mismatches 22;	Indels 1;	Gaps 1

QY	1423	ccgaagcgtctcaataggtttcttgttcctgaaagacacagctcatgtctcaacgagctcgc	1482
Db	68716	CTCTGGCTGGCGATCGCAGGGTTCTTGTTCCTGAAGAACATCAGGGCTATGCTCACCGAGTCTCGC	68655
QY	1483	agcatcaatctctagcacacccacgcgcgtctgcgtcatccaggaatcacaccttggct	1542
Db	68656	AGCATCAATCTCTAGCACACCCAGCGGCTGCTGCGCCATCCAGGAATACACCCCTTTGCT	68599
QY	1543	acgcccctcatgtcaggaagttgttctggcatcacagctaaactagcgcgtctgccttgggtga	16022
Db	68596	ACGCCCCATGTGAGAGAGTGTGTTCCGCAACAGGTACTAGCGCGTCCCTTGGGTGA	68533
QY	1603	caacctcgagagagcagccagaccacagagccctctgagccgcacatcccgggccaagacag	16622
Db	68536	CACCTCGAGAGGACACCCAGACCCAGAGCCCTCTAGCGCGCACATCCGGGCCAAGACAG	68477
QY	1663	atggaacatgcgcgaagggccgacaatgacctgtgtgcctcttccttagggaaatccctgct	17222
Db	68476	ATGGAACTGCGCCAGAGAGCCGGAATAAGCCCTGCTGTGCGCTCTCCCTTAGGGAAATTCCTGCT	68411
QY	1723	atgacagcgtgcctatagattccctcaggaaggaatggtgtgtcccccacaccccagafttcaagtc	1782
Db	68416	ATGACAGCGTGCCTAGATTCCTCTACGAAGAGACATGTGGTCCCCCACCCCCAGTTCAATC	68355
QY	1783	tgtagggagtgaaagccacacgaactcttcaagtgtggaagtgcaactgacactgtatgtcagaacc	18422
Db	68355	TGTAGGAGATGTAAAGCCACAGACTCTTACGTGGAGAGTGCATGTACCTGTAGGTAGTACGAGAC	68229
QY	1843	atacaggaagccaaggttgcaccttctccctttaaagggccctgtgtctcgggagaagaatccc	19022

Db	68296	ATCAGAGAGGCAAGTTCGCCCTTTCCTTTAAAGGCCCTGTGGTCTCGGGAGAAATCC	68237
Qy	1903	tcagatcccacctaagtgctcaaggtgtggaaggacaagcgacaagatagagcatct	1962
Db	68236	TCAGATCCCACTAAAGTGTCAAGGTGTGGAAGGAGACCAAGGACCAAGATGGCCACTCT	68177
Qy	1963	ggggtctctgcccacatccacagtttgtgtctccgtgagctctttcaattgcaacct	2022
Db	68176	GGGGCTAATGCCAATACCCACAGTTTGTGGCTTCCGTGAATCTTTTCATGTGCTACTCT	68117
Qy	2023	aaatgctctgtctcccaattcccaactcgltccctccctctctccagctgtcttgggct	2082
Db	68116	AATATCTCTGTCTCCACTTCCACTTCGTTCCCTCCCTCTCCGAGCTGCTTGTGGGCT	68057
Qy	2083	caagcctgtactcatctggcagagtgctagatctgtgggaagtcctctaagagatgag	2142
Db	68056	CCAGCCCTGTACTCTATCGGACGGCCACAGAGATCTGTGGAGTCTTTAAGAGATGAG	67997
Qy	2143	aagccagaagagcccttcacacaattgtccagaagcttgqacatgacccatcccgacatca	2202
Db	67996	AAGCCAGAGGCGCTCCACCAATGTCAAGACTTGCAATGCACTTCGCGGCACATCA	67937
Qy	2203	tctctgtctctgcacatccatttgaacacattatgaacactgtataatgtagctgtctg	2252
Db	67936	TTCTGTGTCTCTGCATTCATTTGAACACATTAATTAAGCACCGATTAATGATAGCTGTG	67877
Qy	2263	tgggtatatacgcatttgtctccagatatagatctcagctcagctacaggttatagttaaaaa	2322
Db	67876	TGGGTATATACCATATGATCTCGATATAGATCTTAGCTCACAGGTTTATATGTTAAAAA	67817
Qy	2323	aacaaacagaaacacaaacaaatttggatcnaaaggagaaatgtaaagtacgaaaagcag	2382
Db	67816	AACAAACAGAAACACAAACGATTTGGATCAAAAGAG- AATATGATTAAGTGAACAAAGCAG	67758
Qy	2383	cacaaggaaattccctgtgtgtatgtctgagctgtgatgacgagcctgtgacccaagt	2442
Db	67757	CACAAGGAATTTCCCTGTGTGGATGTGATGCTGATGGCGGCACTGGATACCCAAGTG	67698
Qy	2443	aaggtcccgagacatgtgctctgaagagagcaaaagcacaactgcagctgtgaatgct	2502
Db	67697	AAGGTCCCAAGACATGATGCTGTATAGAGCAAGGCAACAACTGACGTGTGATGCGT	67638
Qy	2503	gtgtgtatcttggctgtgaagtagctgttttcgcaactgtatggccctgtgttctctg	2562
Db	67637	GTTGTGTGATTTGGTGTAGGTATAGTCTGTGTGGCACTGATGGGGCTGGGTTGTCTCG	67578
Qy	2563	gggtcgtgaatgttgtgtatgtcctctgtgacaaggtcagctgtgacaatcagtttaacacac	2622
Db	67577	GGGCTGGAATGCTGGGTATGCTCTGTGACAAAGCTACGCTGACATCACTTAAACACAC	67518
Qy	2623	ggagaagaacacattacatgcacctataattctgtgtacacatcbatctcaagctaa	2682
Db	67517	GGAGAAGAACATTATACATGACACCTTAATTTCTGTGTCACATCATTTCTCAAGCTTA	67458
Qy	2683	aaggtatgaagtgcctgcctgtttatatgccaactgttaagaaaaatttttttgcatt	2742
Db	67457	AGGGTATGAAAGTCCCTGCTTGTATTATAGCCACTGTGTAGTAAATATTTTTCGATT	67398
Qy	2743	tcacaattatacttatataaagcatccacacccaaagactagtttgggaaatgtag	2802
Db	67397	TCCAAATTAATACCTTATATTAATTAAGGCAATCCACACTTAAGACATGATTTGGGAAATGTAG	67338
Qy	2803	ccctgggttaatgttcaaatcaaggcaaaagaaatlaaatatgtactcttggctaaaaa	2862
Db	67337	CCCTGGGTTAATGTCAATTAACAGCAAAAGGAATTAATATGTACTTTGGCTAGAGG	67278
Qy	2863	aaaaaa 2868	
Db	67277	GGTAAA 67272	

LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM
AF182217.1	1895 bp	AF182217.1	GI:5853353			European rabbit.
						Oryctolagus cuniculus
						Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi
						Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
						1 (bases 1 to 1895)
REFERENCE						Savas, U., Wester, M.R., Griffin, R.J. and Johnson, E.F.
AUTHORS						Rabbit pregnancy X receptor is activated by rifampicin
JOURNAL						Drug Metab. Dispos. 28 (5), 529-537 (2000)
MEDLINE						20235257.
REFERENCE						2 (bases 1 to 1895)
AUTHORS						Savas, U., Wester, M.R., Griffin, R.J. and Johnson, E.F.
TITLE						Direct Submision
JOURNAL						Submitted (01-SEP-1999) Division of Biochemistry, The Scripps
						Research Institute, 10550 North Torrey Pines Rd., La Jolla, CA
FEATURES						1. location/Qualifiers
						1895

```

source
1..1095
/organism="Oryctolagus cuniculus"
/strain="bred New Zealand white rabbit"
/db_xref="taxon:9986"
/tissue_type="kidney"
1..1895
/gene="NR112"
/note="PXR"
257..1492
257..1492

```

```

BASE COUNT
ORIGIN
434 a      556 c      555 g      370 t
      /gene="nrlz"
      /note="PKR; nuclear receptor subfamily 1 group 1 member 2"
      /zinc-finger protein; transcription factor"
      /codon_start=1
      /product="pregnane x receptor"
      /protein_id="ABD54426.1"
      /db_xref="gi:58535354"
      /translation="MGKPRITISADEEGPQTCNVCGDKANGHFNVLICEGCKGFFRI
      /YKRNARLACPRKACCEITRTTRKQOCARLRKCEGSKKMKMSDVAVAQARAL
      KKKRRPMAPQPGMOGLGEORMIIELDADOMKFDITTESHFKRFLPEVLGSGCC
      TPEISQALTEBEGKMRIOBELCTMLSTLQLEGGESWVNPYPADRSKKLFSL
      HLAIDNSTYFKGILNFAKVAISTYFDDLTDEQIISLKGATLELCIAFNVPNAEGT
      EGGRIQCYCEDPEGGFOQLADVPLKRFHMYLKKQLQKREYLVQMSISLPSDREGV
      GGGVDOQLFAFALITKAVIECSRPQPTFRFLFKLMAVLTETRTINAAQTORLLRTI
      DHNRATPLMRLESTTDD"

```

Query Match	37.9%	Score 1100.4;	DB 3;	Length 1895;
Best Local Similarity	82.4%	Pred No. 3e-216;		
Matches 1288;	Conservative	0;	Mismatches 271;	Indels 5; Gaps 2

QY	116	acgaagaagaataactatactactctgtccctgcgaacaaggcagcgctcttgytaagcta	175
DB	29	AGCGAGAGACCGGCTATTCCTCCGCTCTTTCGACAGGATTACTCATGAGCACTGCTG	88
QY	176	ctccttgatcatcttccttgaccggaatgttcaaaigtgaccccaaggagaaatcggag	235
DB	89	TTC--TGATTATATCTTTGCTCCTGATCCTCAAAATGACCCCAAGGAGGAAGTCACTGAG	146
QY	236	caaaagaacttccaccaagaagcagtcaccaaggagcccggaagaacaaacttgaaagtgcacca	295
DB	147	CCAGGAAGTTTCTGCTCCAAAGCAGGCCCAAGAGGCCCAAAACCAACTGGAGGTCATCTG	206
QY	296	aagaagaagtgaaccactgcgtgactttgtacactgtgagagacacagagctgttccctgaa	355
DB	207	AAGGCAGTGGGAAAACAAGTCGGCCCTTGGCGGCTGTATTAACAACATCTCATGGGTGAA	266
QY	356	agcccaagtgtaaacagacagatcgagaaagtcggaagtccccaaatcgcgtgtatgtggg	415
DB	267	AGCCCAACCATATAGTGCAGATGAGG---GGAAGGGCCCAAGCTTGGCGGTGATCCGGGG	323

QY 416 acaagagccactggtatcattcaatgctcgtgaaagatgcaagggtcttttca 475
 Db 324 ACAAGGCAATGGTTTACCACTTCAACGCTCTGACATGTGAAGCTCAAGGGCTTTTAA 383
 QY 476 ggaagggcactgaaacgcaacgctgagtggtcccttcggaaggcgcttcgaga 535
 Db 384 GGAAGGACCGTGAACCGCAACGGCCGCTGAGATGCCCTTCCGGAAGGGCCCTGGGAGA 443
 QY 536 tcaacgggaagaccgagcagatgcccaggtcgcgcctgcgaagtccctggaagcg 595
 Db 444 TCACCGGAGAACCCGCGCCAGCTGACAGGCTGCGGCTGCGCAAGTCTGTGAAGAGG 503
 QY 596 gcatgaaagagagatgatactgctcgaagagcgctggaagagagcgagccttata 655
 Db 504 GCATGAAGAGAGATGATCATGTGCTGCTGCGACAGACGCGGCTGTGATCA 563
 QY 656 agcgaagaaagatgaaacgagacagcactgagcagtgagtgagtgagtgagtgag 715
 Db 564 AGAGGAAGAAAGAGAGAGAGATGAGGCTCAGCCGCGGGAATGCAAGGGCTGACGGAG 623
 QY 716 agcagcgagatgatactgagagagcgtgagcgtcagaatgaaaccttgaactac 775
 Db 624 AGCAGCGAGATGATATGAGAGAGCTGATGAGCTCAGATGAAACCTTGAACACACCT 683
 QY 776 tctcccaattcaagaaattccggtcgcgcaggggtgcttgcaagtgctgcaggttcag 835
 Db 684 TCTCCCATTTCAAGAAATTTTGGGCTCCGAGAGTGTCTGGCACTGTGTGATTCAG 743
 QY 836 agctctgcagagccactcagagagagagagcgtcagagtgagtgagtgagtgagtg 895
 Db 744 AATCTCTGCAAGGCTTACAGAGAGAGAGAGAGCTGAGCTGAGAGAGATCAAGAGAGC 803
 QY 896 tgtgtctcttgaagagctcctcgtcagcgtcgagggagagagagtgctgagtgag 955
 Db 804 TGGGCAACATGAGAGCTCTCTGCTGACAGCTGAGGAGAGAGAGAGAGAGAGAGAG 863
 QY 956 aaccggcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1015
 Db 864 CGCCCGCAAGTGAACAG 923
 QY 1016 tgcac 1075
 Db 924 TGTCAACATCAATGTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 983
 QY 1076 acttgcac 1135
 Db 984 ACTTGCACATCAAG 1043
 QY 1136 tggatgac 1195
 Db 1044 TGAATTCACACACAGTGTTCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1103
 QY 1196 actgcttggaaagac 1255
 Db 1104 ACTGCTGGAAAGACCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1163
 QY 1256 tccac 1315
 Db 1164 TCCACTACATGCTGAAG 1223
 QY 1316 tctccgcttctcccgagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1375
 Db 1224 TCTCCCTTTCTCCCAAGACCGCCAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1283
 QY 1376 aggaagcaatcgcacatcactcgaagtcctacatgaaacacacacacacacacacacac 1435
 Db 1284 AGGAGGATTTTGGCATCACCTGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1343
 QY 1436 ataggtcttcttctcgaagacatcagcgtcgaacacacacacacacacacacacac 1495
 Db 1344 ACCGCTCTCTGCTGAG 1403
 QY 1496 agcaac 1555

Db 1404 AGCAGACACAGCGGCTGTGCGGATCCAGAGACACACACACCTTGGCACCCCTCATGC 1463
 QY 1556 aggaattgctgacacacagatgctgagagagtgctgacctgagagag 1615
 Db 1464 GGAAGCTTTAGACACACACATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1523
 QY 1616 cagcagac 1675
 Db 1524 CAGCAGGCGCCAGAGCGCTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1583
 QY 1676 agag 1679
 Db 1584 TGTG 1587
 RESULT 8
 LOCUS AF031814 1709 bp mRNA ROD 07-FEB-1998
 DEFINITION Mus musculus pregnane X receptor mRNA, complete cds.
 ACCESSION AF031814
 VERSION AF031814.1 GI:2852328
 KEYWORDS
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 1709)
 AUTHORS Kilewer, S.A., Moore, J.T., Wade, L., Staudinger, J.L., Watson, M.A.,
 Jones, S.A., McKee, D.D., Oliver, B.B., Willson, T.M.,
 Zetterstrom, R.H., Perlmann, T. and Lehmann, J.M.
 An orphan nuclear receptor activated by pregnanes defines a novel
 steroid signaling pathway
 Cell 92 (1), 73-82 (1998)
 JOURNAL MEDLINE 98149345
 REFERENCE 2 (bases 1 to 1709)
 AUTHORS Kilewer, S.A., Moore, J.T., Wade, L., Staudinger, J.L., Watson, M.A.,
 Jones, S.A., McKee, D.D., Oliver, B.B., Willson, T.M.,
 Zetterstrom, R.H., Perlmann, T. and Lehmann, J.M.
 Direct Submission
 TITLE Submitted (29-OCT-1997) Department of Molecular Endocrinology,
 Glaxo Wellcome, 5 Moore Drive, RTP, NC 27709, USA
 FEATURES
 source
 1. 1709
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /tissue_type="liver"
 CDS
 151..1446
 /function="steroid signaling"
 /note="pregnane-activated nuclear receptor"
 /codon_start=1
 /product="pregnane X receptor"
 /protein_id="A039964.1"
 /db_xref="GI:2852329"
 /translation="MRPESWSRVGLVCEADSLAEPINVEEDGGLQICRVCGDK
 ANQYHNVTCEGKCFPRAMKRNRLCPKPGCTETTRKRCOCARAKLES
 GSKKEMIDDAVEORALIKRRKRIKAPPGGGGLLEEOALIOELMDOMOTFD
 TTSHEKDRPLAVHSGCELPEFLQASLEEDPAWMSOIMKRVPMKLSIORGDS
 IMWVOPSSDKGEIRPLPLHADVSTYFKVINYPAKYSIFRDLPIEDQSLKGA
 TEMCLRNINMEDITGTEWEGRLAYCEDENGGKQKLLDPLMKFOMKQLQHK
 EEVLMQALISFSPDRGVORSVVDQLOERALTAKAIECSRPYPAHRLFELKIMA
 VLELRISINAOOTQOLRLRIDSHPFATPLMOELFSSTDG"
 BASE COUNT 449 a 444 c 468 g 348 t
 ORIGIN
 Query Match 34.7%; Score 1006.8; DB 11; Length 1709;
 Best Local Similarity 80.4%; Pred. No. 5.5e-197;
 Matches 1217; Conservative 0; Mismatches 292; Indels 5; Gaps 3;
 QY 179 ctatgatgactcttgacacagagatgctcaaatgagacccacagagagagagtcgagca 238
 Db 44 CTGGGCAAGAACCATGCTTCTTCAAGGTGAGACCCCAAGGGGAATTCACACAA 103

QY 239 agaactaccacacagcgtcccaagagcccaagcaaaccttgagtgagaccacaag 238
 Db AGCGATGGCCCCCAGACAGCTAGGACACAGATGTAACCTAGAGAGAGCTAGG 163
 QY 229 aaactgtgaacacgtctactgtacacgttgagacacagagctcgttccttgaaagc 358
 Db AGAGCTGAGAGCCGAGTGGCTTGTACAGTGTGAAGAAGCAAGACTCTGCTGGAAGAGC 223
 QY 359 ccagtgatcacgcagatgagagagtcgagagtcaccaatctgcgtgtatgtggagca 418
 Db CCA---TCAGAGTAG 280
 QY 419 aggcacacgtgtatcactcactcactcactcactcactcactcactcactcactcactc 478
 Db AGGCAATAGGCTACCTCAATGATGATGATGATGATGATGATGATGATGATGATGATG 340
 QY 479 gggccatgaaac 538
 Db GGGCCATGAAAGCAATGTCGGGGTGAAGTCCCTTCGGAAGGAAAGCTGGAGATCA 400
 QY 539 cccggaagaccccgagacagtgccagcctgcgtgcgtgcgtgcgtgcgtgcgtgcgtgc 598
 Db CCGGAAAGACAG 460
 QY 559 tgaagaaagagatgatactcctcagcagagcgttgagagagagagagagagagagagag 658
 Db TGAAGAAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 520
 QY 659 ggaagaaagagatgatactcctcagcagagcgttgagagagagagagagagagagagag 718
 Db GGAAGAAAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 580
 QY 719 agcagatgatgacag 778
 Db AGCAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 640
 QY 779 cccattcaagagatgatactcctcagcagagcgttgagagagagagagagagagagagag 838
 Db CCACTTCAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 700
 QY 839 ctctcagagcccgagacagtgccagcctgcgtgcgtgcgtgcgtgcgtgcgtgcgtgc 898
 Db TTCTCAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 760
 QY 899 gctcttgaagagatgatactcctcagcagagcgttgagagagagagagagagagagagag 958
 Db TTCCATGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 820
 QY 959 ccccgagacagtgag 1018
 Db CCCCTTCAAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 880
 QY 1019 caactacatgttcaaaagagacatcagccttgccaaagtcactcactcactcagagact 1078
 Db CAACCTACATGTTCAAG 940
 QY 1079 tggccatcagagacagagatcctcgtcgtgaagagagagagagagagagagagagagag 1138
 Db TGGCTATTGAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1000
 QY 1139 gattcaacacagtgatcaaacgagagacagtgagagagagagagagagagagagagagag 1198
 Db GATTCAACAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1060
 QY 1199 gcttgagagacagtgag 1258
 Db GCTTGAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1120
 QY 1259 actacatgttgag 1318
 Db ACTGATGCTGAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1180

QY 1319 cctctctccacagacacccagagtgctgacagacacgcgtgtgtgagcaccagtgcaag 1378
 Db CCCTCTCTCCACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1240
 QY 1379 agcaatcgccttactcactgaaagtcactcactcactcactcactcactcactcactcact 1438
 Db AGAGGTTTCCCTTCAACCCGAG 1300
 QY 1439 gttctctgtctgagagacatgagtgacagagagagagagagagagagagagagagagag 1498
 Db GGTCTGTCTGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1360
 QY 1499 acaccagagtgctgctgac 1558
 Db AACCCAGAGAGTGTGCGGATTCAG 1420
 QY 1559 agttgttgacacagagtgacag 1618
 Db AGTTATTAGCAGACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1480
 QY 1619 ccagaccag 1678
 Db CTAGACCCAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1538
 QY 1679 gccgacacagtgccct 1692
 Db GCCACACAGCCCT 1552

RESULT 9

AF151377 1755 bp mRNA ROD 05-AUG-1999
 LOCUS Rattus norvegicus pregnane X receptor mRNA, complete cds.
 DEFINITION AF151377
 ACCESSION AF151377
 VERSION AF151377.1 GI:5702232
 KEYWORDS Norway rat.
 SOURCE Rattus norvegicus
 ORGANISM Rattus norvegicus; Chordata; Craniata; Vertebrata; Euteleostomi; Eukaryota; Metazoa; Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Rattus.

REFERENCE 1 (bases 1 to 1755)
 AUTHORS Zhang, H., Leculysse, E., Liu, L., Hu, M., Matoney, L., Zhu, W. and Yan, B.
 TITLE Rat Pregnane X Receptor: Molecular Cloning, Tissue Distribution, and Xenobiotic Regulation
 JOURNAL Archive of Biochemistry and Biophysics 368 (1), 14-22 (1999)
 PUBMED 10415106

REFERENCE 2 (bases 1 to 1755)
 AUTHORS Zhang, H., Leculysse, E., Liu, L., Hu, M., Matoney, L., Zhu, W. and Yan, B.
 TITLE Direct Submission
 JOURNAL Submitted (13-MAY-1999) Biomedical Sciences, University of Rhode Island, 41 Lower College Road, Kingston, RI 02881, USA
 FEATURES location/Qualifiers

CDS
 1..1755 /organism="Rattus norvegicus"
 /strain="Sprague-Dawley"
 /db_xref="taxon:10116"
 218..1513 /codon_start=1
 /product="pregnane X receptor"
 /protein_id="AND47214.1"
 /db_xref="GI:5702232"

BASE COUNT 458 a 468 c 480 g 349 t
 ORIGIN
 VTEERSINAOOTQOLLRIODTHPAPATIMOELEFSTDG"


```

RESULT 10
LOCUS AF188476
DEFINITION Oryctolagus cuniculus pregnane X receptor (PXR) mRNA, complete cds
ACCESSION AF188476
VERSION AF188476.1 GI:6901679
KEYWORDS
SOURCE European rabbit.
ORGANISM Oryctolagus cuniculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
REFERENCE 1. (bases 1 to 1601)
Jones, S.A., Moore, L.B., Shenk, J.L., Wisely, G.B., Hamilton, G.A.,
McKee, D.D., Tomkinson, N.C., LeCluyse, E.L., Lambert, M.H.,
Willson, T.M., Kilewer, S.A. and Moore, J.T.
The pregnane X receptor: a promiscuous xenobiotic receptor that has
diverged during evolution
Mol. Endocrinol. 14 (1), 27-39 (2000)
MEDLINE 20092326
PUBMED 10628245
2. (bases 1 to 1601)
Jones, S.A., Moore, L.B., Shenk, J.L., Wisely, G.B., Hamilton, G.A.,
McKee, D.D., Tomkinson, N.C., LeCluyse, E.L., Lambert, M.H.,
Willson, T.M., Kilewer, S.A. and Moore, J.T.
Direct Submission
Submitted (20-SEP-1999) Molecular Endocrinology, Glaxo Wellcome, 5
Moore Drive, RTP, NC 27709, USA
FEATURES
source
1. 1601
location/qualifiers
/organism="Oryctolagus cuniculus"
/db_xref="taxon:9986"
1. 1601
gene

```

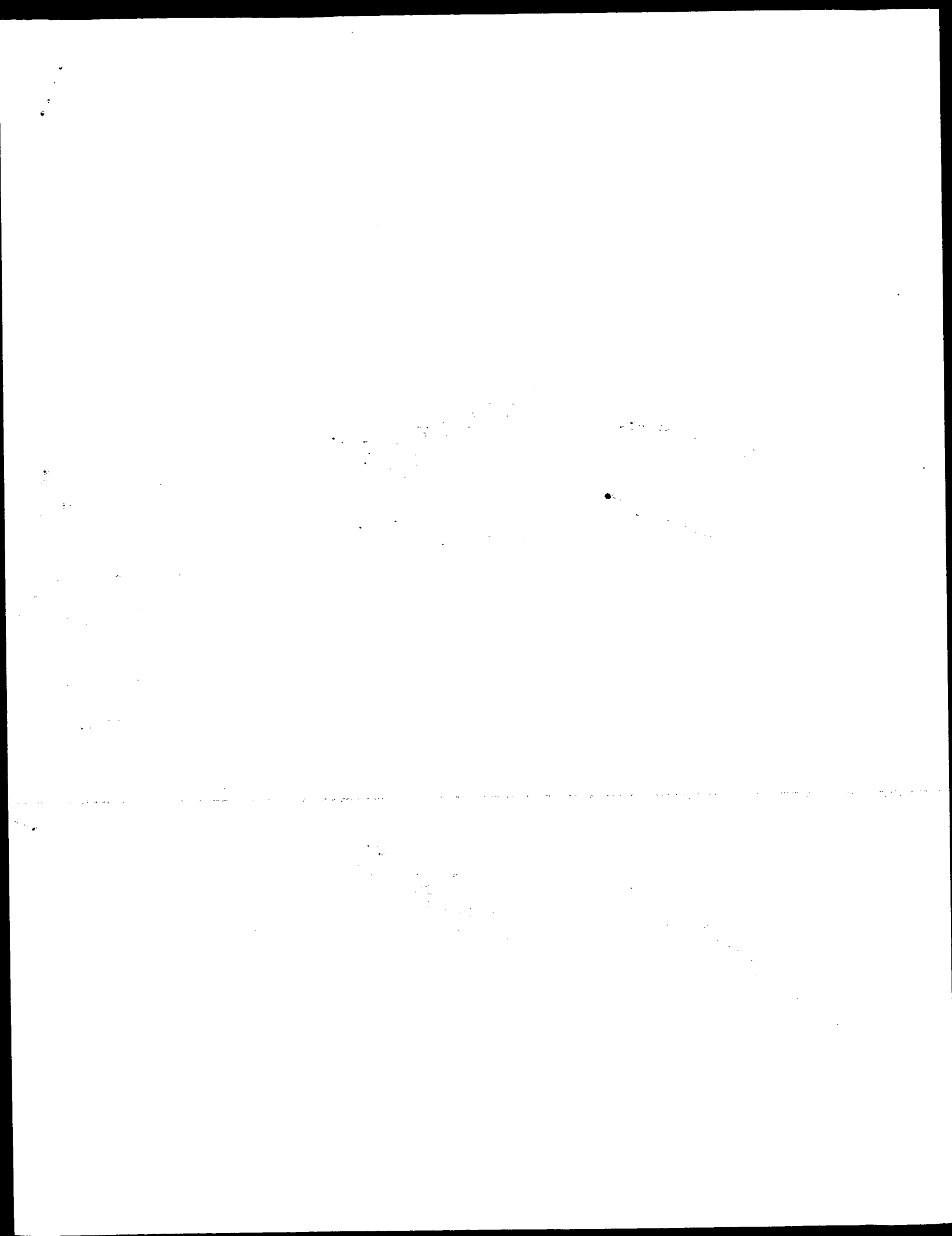

Db 449 CATCAGCATCCTCATCGCCGGCACAACGCACCTTCGACTCCAGCTTCTCCAGTTCCA 508


```

Oy 1441 ttcttgcctgaagatcatgctatctcaccgagctccgcagcatcaatgctcagcac 1500
Db 1315 CTGCTGTACGCCAAGATGATCCAGAGCTGGCGACCTGGCAGCCTCAACGAGAGCAC 1374
Oy 1501 acccagcgagctgctgagcatccagagacataacaccccttctagcgcctcatgcaagag 1560
Db 1375 TCCAAAGCAGTACCGCTGCTCTCTTCACAGCCGAGCACAGCATGCGAGCTCAGCCTCTG 1434
Oy 1561 ttgttcggcatcacaggtacg 1581
Db 1435 GTGCTCGAGGTCTTCGGCAAC 1455

```

Search completed: February 18, 2001, 16:45:23
 Job time: 10903 sec



GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 18, 2001, 11:21:51 ; Search time 67.2 Seconds

(without alignments)
756,968 Million cell updates/sec

Title: US-09-143-828-2

Perfect score: 2290

Sequence: 1 MEYRKPESNHADEHCEPT.....QDIHPATPLMQELFGITGS 434

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues

Total number of hits satisfying chosen parameters: 374700

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: SP archaea:.*
2: SP bacteria:.*
3: SP fungi:.*
4: SP human:.*
5: SP_invertebrate:.*
6: SP_mammal:.*
7: SP_mmc:.*
8: SP_organelle:.*
9: SP_phage:.*
10: SP_plant:.*
11: SP_protent:.*
12: SP_virus:.*
13: SP_vertebrate:.*
14: SP_unclassified:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2287	99.9	473	4	Q9UNW4
2	2281	99.6	434	4	Q9UJ27
3	2278	99.5	457	4	Q9UJ26
4	2061.5	90.0	397	4	Q9UJ24
5	2058.5	89.9	420	4	Q9UJ23
6	1973	88.9	379	4	Q9UJ25
7	1814.5	79.2	411	6	Q9UJ22
8	1753.5	76.9	431	11	Q9UJ21
9	1753.5	76.6	342	4	Q9UJ20
10	991.5	43.3	386	13	Q91B39
11	817	33.7	420	13	Q91B38
12	792	34.6	453	13	Q91B37
13	790.5	34.5	425	13	Q91B36
14	729	31.8	358	11	Q90U51
15	504.5	23.0	445	11	Q90U50
16	498	21.7	455	4	Q92970
17	463.5	20.5	472	4	Q92969
18	459	20.0	397	5	Q76246
19	457	20.0	386	13	Q9PVE4

20	456	19.9	469	11	Q62735	Q62735	rattus norv
21	451.5	19.7	484	11	Q60641	Q60641	mus musc
22	450	19.7	416	13	Q9W6N4	Q9W6N4	hippoglossu
23	447	19.5	673	5	Q76827	Q76827	ceratilis c
24	447	19.5	680	5	Q90314	Q90314	aedes albop
25	445	19.4	456	4	Q13986	Q13986	homo sapien
26	442	19.3	454	13	Q9W524	Q9W524	fugu rubrip
27	434	19.0	447	13	Q9W523	Q9W523	fugu rubrip
28	432	18.9	881	5	Q9V9K8	Q9V9K8	drosophila
29	431.5	18.8	457	13	Q90272	Q90272	brachydanio
30	431.5	18.8	541	5	Q97095	Q97095	locusta mig
31	427	18.6	494	13	Q91391	Q91391	brachydanio
32	426.5	18.6	441	5	Q02035	Q02035	tenebrio mo
33	425	18.6	455	13	Q9W6B3	Q9W6B3	coturnix co
34	424	18.5	459	11	Q90WJ1	Q90WJ1	rattus norv
35	423.5	18.5	452	13	Q91155	Q91155	notophthalm
36	423	18.5	444	13	Q90271	Q90271	brachydanio
37	422	18.4	416	13	Q9W785	Q9W785	salmo salar
38	422	18.4	448	13	Q92019	Q92019	salmo salar
39	421.5	18.4	401	13	Q90382	Q90382	calina mos
40	418	18.3	410	6	Q97715	Q97715	sus scrofa
41	416	18.2	453	13	Q90967	Q90967	gallus gall
42	416	18.2	460	13	Q90966	Q90966	gallus gall
43	416	18.2	462	11	P97513	P97513	mus spretus
44	414.5	18.1	549	13	Q90970	Q90970	gallus gall
45	413	18.0	416	4	Q9P300	Q9P300	homo sapien

ALIGNMENTS

RESULT	ID	PRELIMINARY	PRT	473 AA
1	Q9UNW4			
AC	Q9UNW4			
DT	01-MAY-2000 (TREMBL)	13, Created		
DT	01-MAY-2000 (TREMBL)	13, Last sequence update		
DT	01-OCT-2000 (TREMBL)	15, Last annotation update		
DE	ORPHAN NUCLEAR RECEPTOR.			
GN	PAR2.			
OS	Homo sapiens (human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Carnivora; Homiidae; Homo.			
OX	NCBI_Taxid:9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE-LIVER;			
RX	MEDLINE:98445350; PubMed:9770465;			
RA	Bertilsson G., Heidrich J., Svensson K., Asman M., Jendeborg L.,			
RA	Sydow-Bachman M., Ohlsson R., Postlund H., Blomquist P.,			
RA	Berkenstam A.,			
RT	"Identification of a human nuclear receptor defines a new signaling			
RT	pathway for CYP3A induction."			
RL	Proc. Natl. Acad. Sci. U.S.A. 95:12208-12213(1998).			
DR	EMBL: AF084644; AAC64557.1;			
DR	HSSP: P10826; IIRA.			
DR	INTERPRO: IPR000324;			
DR	INTERPRO: IPR000536;			
DR	INTERPRO: IPR001628;			
DR	INTERPRO: IPR001723;			
DR	PFAM: PF00104; hormone_rec; 1.			
DR	PFAM: PF00105; zf-C4; 1.			
DR	PRINTS: PR00047; STROLDINGER.			
DR	PRINTS: PR00350; VITAMINDR.			
DR	PRINTS: PR00396; STRORHOMER.			
DR	PROSITE: PS00031; NUCLEAR_RECEPTOR; 1.			
KW	Receptor.			
SO	SEQUENCE	473 AA; 53899 MW; 3049D23C9B903016 CRC64;		

191958
X ref good H

Query Match 99.9%; Score 2287; DB 4; Length 473;
Best Local Similarity 99.8%; Pred. No. 6.9e-185;
Matches 433; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEVPRKSNHADPVHCEDETESPGKPSVNADEVEGPOICRVCGDKATGYHFNMTCEG 60
 DB 40 LEVPRKSNHADPVHCEDETESPGKPSVNADEVEGPOICRVCGDKATGYHFNMTCEG 99
 QY 61 CKGFPRAMRNARLRCPPFKGACETTRKTRRQCAQLRKCLSGMKKEMIMSDAEVE 120
 DB 100 CKGFPRAMRNARLRCPPFKGACETTRKTRRQCAQLRKCLSGMKKEMIMSDAEVE 159
 QY 121 RRALKRKRSERTGTPGLVQGLTEEOGRMIRELMDAQKRTDPTFSHKRNLPGVLS 180
 DB 160 RRALKRKRSERTGTPGLVQGLTEEOGRMIRELMDAQKRTDPTFSHKRNLPGVLS 219
 QY 181 GCELPESLOAPSRBEAKMSQVARKDLSKVSQQLRGEDGSWMNYPADSGKEIFSL 240
 DB 220 GCELPESLOAPSRBEAKMSQVARKDLSKVSQQLRGEDGSWMNYPADSGKEIFSL 279
 QY 241 PHMADSTYMFKGIISFAKVISYFRDLPIEDQISLKGAFELCOLRNTVNAETGWE 300
 DB 280 PHMADSTYMFKGIISFAKVISYFRDLPIEDQISLKGAFELCOLRNTVNAETGWE 339
 QY 301 CGRLSTCEDYAGGFOQLLEPMKFKHYMLKQLHHEEYVLMQALISLSPDRPGVLOHR 360
 DB 340 CGRLSTCEDYAGGFOQLLEPMKFKHYMLKQLHHEEYVLMQALISLSPDRPGVLOHR 399
 QY 361 VVDQLOEOPAFITLKSYTECNRPQAPARFLFLKIMAMTELRSINQHTORLLRIODIHF 420
 DB 400 VVDQLOEOPAFITLKSYTECNRPQAPARFLFLKIMAMTELRSINQHTORLLRIODIHF 459
 QY 421 ATPMOELFGITGS 434
 DB 460 ATPMOELFGITGS 473

RESULT 2

Q9UJ27 PRELIMINARY: PRT: 434 AA.
 AC Q9UJ27;
 DT 01-MAY-2000 (Tremblrel. 13, Created)
 DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
 DE 01-OCT-2000 (Tremblrel. 13, Last annotation update)
 DE NUCLEAR HORMONE RECEPTOR PRRI-A.
 GN PRRI.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_Taxid-9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=LIVER;
 RA Heard D.J., Holloway J., Hansen C., Tommerup N., Aagaard L.,
 RA Vissing H.;
 RT "Identification of a novel protein isoform of the human nuclear
 RT hormone receptor PKR and localization to chromosome 3q12.1
 RT -13.3.";
 RL Eur. J. Hum. Genet. 0:0-0(0).
 DR EMBL: AJ009936; CAB55489.1.
 DR HSSP: P10826; IHRA.
 DR INTERPRO: IPR000324;
 DR INTERPRO: IPR000536;
 DR INTERPRO: IPR001628;
 DR INTERPRO: IPR001723;
 DR PFAM: PF00104; hormone_rec. 1.
 DR PFAM: PF00105; zf-C4. 1.
 DR PRINTS: PR00047; STROIDFINGER.
 DR PRINTS: PR00350; VITAMINDR.
 DR PRINTS: PR00398; STRDHOMONER.
 DR PROSITE: PS00031; NUCLEAR_RECEPTOR. 1.
 KW Receptor.
 SQ SEQUENCE 434 AA: 49777 MW: 4836DECC3C4C4200 CRC64;

Query Match

99.6%; Score 2281; DB 4; Length 434;

Best Local Similarity 99.5%; Pred. No. 26-184; Matches 432; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MEVPRKSNHADPVHCEDETESPGKPSVNADEVEGPOICRVCGDKATGYHFNMTCEG 60
 DB 1 MEVPRKSNHADPVHCEDETESPGKPSVNADEVEGPOICRVCGDKATGYHFNMTCEG 60
 QY 61 CKGFPRAMRNARLRCPPFKGACETTRKTRRQCAQLRKCLSGMKKEMIMSDAEVE 120
 DB 61 CKGFPRAMRNARLRCPPFKGACETTRKTRRQCAQLRKCLSGMKKEMIMSDAEVE 120
 QY 121 RRALKRKRSERTGTPGLVQGLTEEOGRMIRELMDAQKRTDPTFSHKRNLPGVLS 180
 DB 121 RRALKRKRSERTGTPGLVQGLTEEOGRMIRELMDAQKRTDPTFSHKRNLPGVLS 180
 QY 181 GCELPESLOAPSRBEAKMSQVARKDLSKVSQQLRGEDGSWMNYPADSGKEIFSL 240
 DB 181 GCELPESLOAPSRBEAKMSQVARKDLSKVSQQLRGEDGSWMNYPADSGKEIFSL 240
 QY 241 PHMADSTYMFKGIISFAKVISYFRDLPIEDQISLKGAFELCOLRNTVNAETGWE 300
 DB 241 PHMADSTYMFKGIISFAKVISYFRDLPIEDQISLKGAFELCOLRNTVNAETGWE 300
 QY 301 CGRLSTCEDYAGGFOQLLEPMKFKHYMLKQLHHEEYVLMQALISLSPDRPGVLOHR 360
 DB 301 CGRLSTCEDYAGGFOQLLEPMKFKHYMLKQLHHEEYVLMQALISLSPDRPGVLOHR 360
 QY 361 VVDQLOEOPAFITLKSYTECNRPQAPARFLFLKIMAMTELRSINQHTORLLRIODIHF 420
 DB 361 VVDQLOEOPAFITLKSYTECNRPQAPARFLFLKIMAMTELRSINQHTORLLRIODIHF 420
 QY 421 ATPMOELFGITGS 434
 DB 421 ATPMOELFGITGS 434

RESULT 3

Q9UJ26 PRELIMINARY: PRT: 457 AA.
 AC Q9UJ26;
 DT 01-MAY-2000 (Tremblrel. 13, Created)
 DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
 DE 01-OCT-2000 (Tremblrel. 15, Last annotation update)
 DE NUCLEAR HORMONE RECEPTOR PRRI-C.
 GN PRRI.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_Taxid-9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=LIVER;
 RA Heard D.J., Holloway J., Hansen C., Tommerup N., Aagaard L.,
 RA Vissing H.;
 RT "Identification of a novel protein isoform of the human nuclear
 RT hormone receptor PKR and localization to chromosome 3q12.1
 RT -13.3.";
 RL Eur. J. Hum. Genet. 0:0-0(0).
 DR EMBL: AJ009936; CAB55490.1.
 DR HSSP: P10826; IHRA.
 DR INTERPRO: IPR000324;
 DR INTERPRO: IPR000536;
 DR INTERPRO: IPR001628;
 DR INTERPRO: IPR001723;
 DR PFAM: PF00104; hormone_rec. 1.
 DR PFAM: PF00105; zf-C4. 1.
 DR PRINTS: PR00047; STROIDFINGER.
 DR PRINTS: PR00350; VITAMINDR.
 DR PRINTS: PR00398; STRDHOMONER.
 DR PROSITE: PS00031; NUCLEAR_RECEPTOR. 1.
 KW Receptor.
 SQ SEQUENCE 457 AA: 52145 MW: 812F8B3545B784E CRC64;

Query Match 99.5%; Score 2278; DB 4; Length 457;
 Best Local Similarity 99.3%; Pred. No. 3.8e-184;
 Matches 431; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 MEVRRKESNNHADVHCDTESVPGKPSVNADEYGGPOICRVCGDKATGYHFNMTCEG 60
 DB 24 LEVRKESNNHADVHCDTESVPGKPSVNADEYGGPOICRVCGDKATGYHFNMTCEG 83
 QY 61 CKGFERRAKNNARLRCPPFRKACETITRTKTRCOACRLKRCLESCKMKEMIMSDAEVEE 120
 DB 84 CKGFERRAKNNARLRCPPFRKACETITRTKTRCOACRLKRCLESCKMKEMIMSDAEVEE 143
 QY 121 RRLIKRKKSERTGTQPLGVGGLTEEQRMIMRELMDAQMKTFTDTSFHKFRPLPGVLS 180
 DB 144 RRLIKRKKSERTGTQPLGVGGLTEEQRMIMRELMDAQMKTFTDTSFHKFRPLPGVLS 203
 QY 181 GCELPESLOAPSRREAAKMSQVRKDLCSLKVSLQLRGDSVWNNKPPADSGGKEIFSL 240
 DB 204 GCELPESLOAPSRREAAKMSQVRKDLCSLKVSLQLRGDSVWNNKPPADSGGKEIFSL 263
 QY 241 PHMADSTYMEKGIISPAKVISYFRDLPIDQISLKGAEELCOLRNTVFNATGTWE 300
 DB 264 PHMADSTYMEKGIISPAKVISYFRDLPIDQISLKGAEELCOLRNTVFNATGTWE 323
 QY 301 CGRLSYCLEDTAGGFOQLLEPMKLFHYMLKKIQLHBEYVLMQATSLFSPDRPGVLOHR 360
 DB 324 CGRLSYCLEDTAGGFOQLLEPMKLFHYMLKKIQLHBEYVLMQATSLFSPDRPGVLOHR 383
 QY 361 VVDLOQOPAITLKSYIECNRPQPAHRELFKIMAMTELRSINAQHTORLRIQDIHPF 420
 DB 384 VVDLOQOPAITLKSYIECNRPQPAHRELFKIMAMTELRSINAQHTORLRIQDIHPF 443
 QY 421 ATPLMQELFGITGS 434
 DB 444 ATPLMQELFGITGS 457

RESULT 4
 ID 090J24 PRELIMINARY; PRT; 397 AA.

AC 090J24; PRELIMINARY; PRT; 397 AA.
 DT 01-MAY-2000 (T-EMBLrel. 13, Created)
 DT 01-MAY-2000 (T-EMBLrel. 13, Last sequence update)
 DT 01-OCT-2000 (T-EMBLrel. 15, Last annotation update)
 DE NUCLEAR HORMONE RECEPTOR PRR2-A.
 GN PRR2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC TISSUE=LIVER;
 RA Heard D.J., Holloway J., Hansen C., Tommerup N., Aagaard L.,
 RA Vissing H.;
 RT Identification of a novel protein isoform of the human nuclear
 RT hormone receptor PKR/SXK and localization to chromosome 3q12.1
 RT -13.3.";
 RL Eur. J. Hum. Genet. 0:0-0(0)
 DR EMBL: AJ009937; CAB5492.1;
 DR HSSP: P10826; IHRA.
 DR INTERPRO: IPR000324;
 DR INTERPRO: IPR000536;
 DR INTERPRO: IPR001628;
 DR INTERPRO: IPR001723;
 DR PFAM: PF00104; hormone_rec; 1.
 DR PRINTS: PR00047; STROIDINGER.
 DR PRINTS: PR00350; VITAMINDR.
 DR PRINTS: PR00398; STRDHOMER.
 DR PRINTS: PR00546; THYROIDHOMR.

DR PROSITE: PS00031; NUCLEAR_RECEPTOR; 1.
 KW Receptor.
 SQ SEQUENCE 397 AA; 45782 MW; 5E2CE0BB10663073 CRC64;

Query Match 90.0%; Score 2061.5; DB 4; Length 397;
 Best Local Similarity 91.0%; Pred. No. 5.8e-166;
 Matches 395; Conservative 0; Mismatches 2; Indels 37; Gaps 1;

QY 1 MEVRRKESNNHADVHCDTESVPGKPSVNADEYGGPOICRVCGDKATGYHFNMTCEG 60
 DB 1 MEVRRKESNNHADVHCDTESVPGKPSVNADEYGGPOICRVCGDKATGYHFNMTCEG 60
 QY 61 CKGFERRAKNNARLRCPPFRKACETITRTKTRCOACRLKRCLESCKMKEMIMSDAEVEE 120
 DB 61 CKGFERRAKNNARLRCPPFRKACETITRTKTRCOACRLKRCLESCKMKEMIMSDAEVEE 120
 QY 121 RRLIKRKKSERTGTQPLGVGGLTEEQRMIMRELMDAQMKTFTDTSFHKFRPLPGVLS 180
 DB 121 RRLIKRKKSERTGTQPLGVGGLTEEQRMIMRELMDAQMKTFTDTSFHKFRPLPGVLS 173
 QY 181 GCELPESLOAPSRREAAKMSQVRKDLCSLKVSLQLRGDSVWNNKPPADSGGKEIFSL 240
 DB 121 RRLIKRKKSERTGTQPLGVGGLTEEQRMIMRELMDAQMKTFTDTSFHKFRPLPGVLS 173
 QY 181 GCELPESLOAPSRREAAKMSQVRKDLCSLKVSLQLRGDSVWNNKPPADSGGKEIFSL 240
 DB 174 -----VSLQLRGDSVWNNKPPADSGGKEIFSL 203
 QY 241 PHMADSTYMEKGIISPAKVISYFRDLPIDQISLKGAEELCOLRNTVFNATGTWE 300
 DB 204 PHMADSTYMEKGIISPAKVISYFRDLPIDQISLKGAEELCOLRNTVFNATGTWE 263
 QY 301 CGRLSYCLEDTAGGFOQLLEPMKLFHYMLKKIQLHBEYVLMQATSLFSPDRPGVLOHR 360
 DB 264 CGRLSYCLEDTAGGFOQLLEPMKLFHYMLKKIQLHBEYVLMQATSLFSPDRPGVLOHR 323
 QY 361 VVDLOQOPAITLKSYIECNRPQPAHRELFKIMAMTELRSINAQHTORLRIQDIHPF 420
 DB 324 VVDLOQOPAITLKSYIECNRPQPAHRELFKIMAMTELRSINAQHTORLRIQDIHPF 383
 QY 421 ATPLMQELFGITGS 434
 DB 384 ATPLMQELFGITGS 397

RESULT 5
 ID 090J23 PRELIMINARY; PRT; 420 AA.

AC 090J23; PRELIMINARY; PRT; 420 AA.
 DT 01-MAY-2000 (T-EMBLrel. 13, Created)
 DT 01-MAY-2000 (T-EMBLrel. 13, Last sequence update)
 DT 01-OCT-2000 (T-EMBLrel. 15, Last annotation update)
 DE NUCLEAR HORMONE RECEPTOR PRR2-C.
 GN PRR2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC TISSUE=LIVER;
 RA Heard D.J., Holloway J., Hansen C., Tommerup N., Aagaard L.,
 RA Vissing H.;
 RT Identification of a novel protein isoform of the human nuclear
 RT hormone receptor PKR/SXK and localization to chromosome 3q12.1
 RT -13.3.";
 RL Eur. J. Hum. Genet. 0:0-0(0)
 DR EMBL: AJ009937; CAB5493.1;
 DR HSSP: P10826; IHRA.
 DR INTERPRO: IPR000324;
 DR INTERPRO: IPR000536;
 DR INTERPRO: IPR001628;
 DR INTERPRO: IPR001723;
 DR INTERPRO: IPR001728;
 DR PFAM: PF00104; hormone_rec; 1.
 DR PRINTS: PR00105; zf-C4; 1.

DR PROTEIN, PS00031, NUCLEAR_RECEPTOR; 1.
DR PROSTATE.
DR PRINTS, PS00036; THYROIDHORM.
DR PRINTS, P800496;
DR PRINTS, P800398; STDHOMOMER.
DR PRINTS, P800047; STROIDEINGER.
DR PRINTS, P800350; VITAMINDR.
SQ SEQUENCE# 420 AA; 48149 MW; CFFD9DEL256859165 CRC64;

Query Match	89.9%	Score 2058.5	-DB 4	Length 420
BestLocal Similarity	90.8%	Pred. No. 1.1e-165		
Matches 39;	Conservative	1;	Mismatches 2;	Indels 37;
				Gaps 1;

0Y	1	MEVVRKEEMNNHADFEHCDEITSEVSGKPSVYNADEVGSPQICRVCQDAKATISIHFNVMITCSG	80
0Y	24	LEVVRKEEMNNHDFEHCDEITSEVSGKPSVYNADEVGSPQICRVCQDAKATISIHFNVMITCSG	83
0Y	61	CKGFFPRAMKRNARLRCPFRKGACETTRKTRPQOACRLKCLSESGKKKEMIMSDAVEE	120
0Y	84	CKGFFPRAMKRNARLRCPFRKGACETTRKTRPQOACRLKCLSESGKKKEMIMSDAVEE	143
0Y	121	RRALTKRRKSKSEKGTQOPRGVQGLTEQOQRMIRLEMDAQMTFDTTSESHFNFRPLPGVLS	180
0Y	144	RRALTKRRKSKSEKGTQOPRGVQGLTEQOQRMIRLEMDAQMTFDTTSESHFNFR-----	196
0Y	181	GCELPESLQAPSRREBAKMSOVRDCLSKVLSLOLRGEGDSVWNYKRPADSGKEIFSL	240
0Y	197	-----VSLQIRGEDSGVWNYKRPADSGKEIFSL	226
0Y	241	PRHADMSTYMEKGISPAKVISYFNDPLTEIOISILKGAPELQOLRFNTVFNAAETGWE	300
0Y	227	PRHADMSTYMEKGISPAKVISYFNDPLTEIOISILKGAPELQOLRFNTVFNAAETGWE	286
0Y	301	CGRLSYCELDTAGCFQOQLLEBPLMFHYMLKRLQLEHEEYVLMQASISLSPDRPGVLOHR	360
0Y	287	CGRLSYCELDTAGCFQOQLLEBPLMFHYMLKRLQLEHEEYVLMQASISLSPDRPGVLOHR	346
0Y	361	VVDQLOEOFAITLKSTIECNRPQPAHRRLFLKINAMLTSLASINAQHTORLLRQIDIHPE	420
0Y	347	VVDQLOEOFAITLKSTIECNRPQPAHRRLFLKINAMLTSLASINAQHTORLLRQIDIHPE	406
0Y	421	ATPLMOELGINGS	434
0Y	407	ATPLMOELGINGS	420

```

RESULT 6
090U25 PRELIMINARY; PRT: 379 AA.
ID 090U25;
AC 090U25;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE NUCLEAR HORMONE RECEPTOR PRR1-B.
GN PRR1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostom
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RA Heard D.J., Holloway J., Hansen C., Tommerup N., Aagaard L.,
RT Vissing H.;
RT Identification of a novel protein isoform of the human nuclear
RT hormone receptor PXR/SXR and localization to chromosome 3q12.1
RT -13.3."
RL Eur. J. Hum. Genet. 0:0-0(0).
DR EMBL: AJ009936; CAB55491.1; -.
DR HSSP: P10826; IHRA.
DR INTERPRO: IPR000324; -.
DR INTERPRO: IPR000536; -.
DR INTERPRO: IPR001628; -.

```

DR INTERPRO: IPR001723; -.
DR PFAM: PF00104; hormone_rec1.
DR PFAM: PF00105; zf-CA_1.
DR PRINTS: PR00047; STEROIDFINGER.
DR PRINTS: PR00350; VITAMINDR.
DR PRINTS: PR00398; STDHORMONER.
KW Receptor.
SQ SEQUENCE 379 AA; 43692 MW; 8F7B7AFA13E45036 CRC64;

Query Match	86.2%	Score 1973;	DB 4;	Length 379;
Best Local Similarity	99.7%	Pred. No. 1.6e-158;		
Matches 378; Conservative	0;	Mismatches 1;	Indels 0;	Gaps 0

```

0Y 56 MTCBCKCFEPRANKRNARLCCPRKACCEITRTTRQCCACRLKLCBESGMKEIMSD 115
Db 1 MTCBCKCFEPRANKRNARLCCPRKACCEITRTTRQCCACRLKLCBESGMKEIMSD 60
0Y 116 EAVEERRALLIRKKTSERTGTPLGVGVLTEBORMMIELMDAQMKTEDTTFSEFKNRLP 175
Db 61 EAVEERRALLIRKKTSERTGTPLGVGVLTEBORMMIELMDAQMKTEDTTFSEFKNRLP 120
0Y 176 GVLSSGCELPISLAPSPSEBAKKSQVRKDLCSIKYSLQIRGEDGSYWNKPPADSGKE 235
Db 121 GVLSSGCELPISLAPSPSEBAKKSQVRKDLCSIKYSLQIRGEDGSYWNKPPADSGKE 180
0Y 236 IFSLLPHMADSYVAFKGIISFAYISYFRDLPLEDQISLLKGAFFCOLRENTYFNAE 295
Db 181 IFSLLPHMADSYVAFKGIISFAYISYFRDLPLEDQISLLKGAFFCOLRENTYFNAE 240
0Y 296 TGTWECGRSLSTLEDTAGGFQOLLLEPMLKTHYMLAKKQLQHEEYVYLMQAIISLSPDRPG 355
Db 241 TGTWECGRSLSTLEDTAGGFQOLLLEPMLKTHYMLAKKQLQHEEYVYLMQAIISLSPDRPG 300
0Y 356 VLOHRYVDOLOBOFKLITLKSYTECNRPQPARBFILTKIMAMLTLSRSINNOHTORLLRIQ 415
Db 301 VLOHRYVDOLOBOFKLITLKSYTECNRPQPARBFILTKIMAMLTLSRSINNOHTORLLRIQ 360
0Y 416 DIHPRATPLMOELFGITGS 434
Db 361 DIHPRATPLMOELFGITGS 379

```

```

RESULT 7
Q9TU02
ID Q9TU02 PRELIMINARY; PRT; 411 AA.
AC Q9TU02;
DT 01-MAY-2000 (TIREMBIrel. 13, Created)
DT 01-MAY-2000 (TIREMBIrel. 13, Last sequence update)
DT 01-OCT-2000 (TIREMBIrel. 15, Last annotation update)
DE PREGNANE X RECEPTOR.
OS NR1I2 OR FXR.
SN Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniala; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN-BRED NEW ZEALAND WHITE RABBIT; TISSUE-KIDNEY;
RA Savas U., Wester M.R., Griffin R.J., Johnson E.F.;
RT "The rabbit pregnane x receptor is activated by rifampicin."
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RN SEQUENCE FROM N.A.
RX MEDLINE=20092326; PubMed=10628745;
RA Jones S.A., Moore L.B., Shenk J.L., Wisely G.B., Hamilton G.A.,
RA McKee D.D., Tomkinson N.C., Lecluyse E.L., Lambert M.H., Willson T.M.,
RA Klierer S.A., Moore J.T.;
RT "The pregnane x receptor: a promiscuous xenobiotic receptor that has
RT diverged during evolution."
RL MOL. ENDOCRINOL. 14:27-39(2000).
DR EMBL; AF182217; AAD54426.1; -.
DR EMBL; AF188476; AAF31165.1; -.

```

AT Mol. Endocrinol. 14: 27-39 (2000).
DR EMBL; AF182217; AAD54426.1; -
DR EMBL; AF188476; AAF31165.1; -

AT Mol. Endocrinol. 14: 27-39 (2000).
DR EMBL; AF182217; AAD54426.1; -
DR EMBL; AF188476; AAF31165.1; -

DR INTERPRO: IPR001628; -
 DR INTERPRO: IPR001723; -
 DR INTERPRO: IPR001728; -
 DR PRAM: PF00104; hormone_rec; 1.
 DR PRAM: PF00104; zt-C4; 1.
 DR PRINTS: PRO0047; STEROIDFINGER.
 DR PRINTS: PRO0350; VITAMINDR.
 DR PRINTS: PRO0398; STEROHORMONER.
 DR PRINTS: PRO0546; THYROIDHORM.
 DR RECEPTOR.
 DR SEQUENCE 342 AA; 39696 MW; 709467C3E9000A3A CRC64;

Query Match 76.6%; Score 1753.5; DB 4; Length 342;
 Best Local Similarity 90.0%; Pred. No. 4.7e-140;
 Matches 341; Conservative 0; Mismatches 1; Indels 37; Gaps 1;

QY 56 MTCGCKGFFRRAMKRNALRCPFRKACETTRTRRQCAQLRKLCSGKMKEMIMSD 115
 DB 1 MTCGCKGFFRRAMKRNALRCPFRKACETTRTRRQCAQLRKLCSGKMKEMIMSD 60
 QY 116 EAVERRALLIRKKSERGTPLGVGITEPQRMIRLMDAOKTDTTFSHFKNRRLP 175
 DB 61 EAVERRALLIRKKSERGTPLGVGITEPQRMIRLMDAOKTDTTFSHFKNR-- 118
 QY 176 GVLSSGCELPESLQAPSHREAAKMSOVRKDLCSLKSLSLQSGEDGSVMNKKPPADSGKE 235
 DB 119 -----VSLQLRGEGGSVMNKKPPADSGKE 143
 QY 236 IFSLLPHMADNSTMFKGISFAVVISYFNDLPIDQISLLKGAAPFLCOLRNTVYNAE 295
 DB 144 IFSLLPHMADNSTMFKGISFAVVISYFNDLPIDQISLLKGAAPFLCOLRNTVYNAE 203
 QY 296 TGTGECGRSLCYLDDTAGGPOQLLEPMKFNHYMLKTLQHEEYVYMOAISLSPDRPG 355
 DB 204 TGTGECGRSLCYLDDTAGGPOQLLEPMKFNHYMLKTLQHEEYVYMOAISLSPDRPG 263
 QY 356 VLOHRVVDQLOEOPATITLSTYECNRPQAPARFLFLKIMAMTELRSINAOHTORLLRIO 415
 DB 264 VLOHRVVDQLOEOPATITLSTYECNRPQAPARFLFLKIMAMTELRSINAOHTORLLRIO 323
 QY 416 DIHPFATPIMOELFGITGS 434
 DB 324 DIHPFATPIMOELFGITGS 342
 RESULT 10
 Q91839 PRELIMINARY; PRT; 386 AA.
 AC Q91839;
 DT 01-NOV-1996 (Tremblrel. 01, Created)
 DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
 DE 01-MAY-2000 (Tremblrel. 13, Last annotation update)
 OS ORPHAN NUCLEAR RECEPTOR OF STEROID/THYROID SUPERFAMILY.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipidae;
 OC Xenopodinae; Xenopus.
 NC NCBL_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94173664; PubMed=8127656;
 RA Smith D.P., Mason C.S., Jones E.A., Old R.W.;
 RT "A novel nuclear receptor superfamily member in Xenopus that
 RT associates with RXR, and shares extensive sequence similarity to the
 RT mammalian vitamin D3 receptor."
 RL Nucleic Acids Res. 22:66-71(1994).
 DR EMBL: X75163; CAA53006.1; -
 DR HSSP: P10826; IHRA.
 DR INTERPRO: IPR000536; -
 DR INTERPRO: IPR001628; -
 DR PRAM: PRO0104; hormone_rec; 1.
 DR PRAM: PRO0105; zt-C4; 1.

DR PRINTS: PRO0047; STEROIDFINGER.
 DR PROSITE: PS00031; NUCLEAR_RECEPTOR; UNKNOWN 1.
 DR SEQUENCE 386 AA; 44302 MW; 88046ED219556573 CRC64;

Query Match 43.3%; Score 991.5; DB 13; Length 386;
 Best Local Similarity 49.0%; Pred. No. 1.1e-75;
 Matches 204; Conservative 59; Mismatches 104; Indels 49; Gaps 9;

QY 18 EDTESVGRKPSVNADEYGGPOICRVGDKATGYHFNMTCEGCKGFFRRAMKRNALRCP 77
 DB 14 EEDSASNSCGTGEDEDDGDPKICRACGRATGYHFNMTCEGCKGFFRRAMKRNALRCP 73
 QY 78 PFRKACETTRTRRQCAQLRKLCSGKMKEMIMSEAVERRALLIRK-KSERGTQ 136
 DB 74 PF-QNSCVINKSNRHCQCRKLKCLDYGKRELIMSAVAEQRALLIRKHKTLKLPPT 132
 QY 137 PLGVGITEPQRMIRLMDAOKTDTTFSHFKNRRLPGLVSSGCELPESLQAPSHREA 196
 DB 133 PPGA-SLPEQGHFLQLVGATKTFDFNFTFSKNR-----PIR--- 171
 QY 197 AKMSOVRKDL:SLKVSLSLQSGEDGSVMNKKPPADSGKEIFSLPHMADNSTMFKGIS 256
 DB 172 -----RSSDPT--QEPQATS--SEATLMPHISDLTYIMKIGIS 207
 QY 257 FAKVISYFNDLPIDQISLLKGAAPFLCOLRNTVYNAETGTCGRSLCYLDDT-AGGF 315
 DB 208 FAKMLPFSLSLIDQIALKGSVAESVIRNNTVFNDSITWEGGPTTYTEMFLAGF 267
 QY 316 QQLLEPMKFNHYMLKTLQHEEYVYMOAISLSPDRPGVLOHRVVDQLOEOPATITLKS 375
 DB 268 RQLFLEPLVIRHMRKNTLQSEEVYAMAAISLSPDRPGVCDMEKIKQLQEHATLTKD 327
 QY 376 YIECNR-POPAREFLKIMAMTELRSINAOHTORLLRIDDIHPATPIMOELFG 430
 DB 328 FIDSORPSPQNRLLPYRIMCETLRLTVNDIHSKOLLEIMDIODAPTPLMRVEVG 383

RESULT 11
 Q91B73 PRELIMINARY; PRT; 420 AA.
 AC Q91B73;
 DT 01-OCT-2000 (Tremblrel. 15, Created)
 DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
 DT 01-OCT-2000 (Tremblrel. 15, Last annotation update)
 DE VITAMIN D RECEPTOR A.
 GN VDRA.
 OS Paratichthys olivaceus (Flounder).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorphi; Acanthopterygii; Percomorpha; Pleuronectiformes;
 OC Pleuronectidae; Bochiidae; Paratichthys.
 NC NCBL_TaxID=8253;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=INTESTINE;
 RA Suzuki T., Suzuki N., Srinastava A.S., Kurokawa T.;
 RT "Identification of cDNAs encoding two subtypes of vitamin D receptor
 RT in flounder, Paratichthys olivaceus."
 RL Biochem. Biophys. Res. Commun. 270:40-45(2000).
 DR EMBL: AB037674; BAA95016.1; -
 DR RECEPTOR.
 DR SEQUENCE 420 AA; 47486 MW; 038FBF00D4F38067 CRC64;

Query Match 35.7%; Score 817; DB 13; Length 420;
 Best Local Similarity 43.1%; Pred. No. 6.3e-61;
 Matches 179; Conservative 70; Mismatches 144; Indels 22; Gaps 12;
 QY 29 VNADL-EVGGPOICRVGDKATGYHFNMTCEGCKGFFRRAMKRNALRCPFRKACETIT 87
 DB 11 VGPDEFDRNAPRICGVGDKATGFHFNMTCEGCKGFFRRAMKRNALRCPFRKACETIT 69

[illegible][illegible]

```

Db 234 VSXIOKVGAFAMIPGFELETAEDQIALKSSAIEVIMLRNSQNFEDMSWCGADPF 293
OY 305 SYCLEL-TAGGPOQLLEPMKFKHYMLKTILOHEEYVMOAISLSPDRPGVLOHRYVD 363
Db 294 KVOISDVTAGHTLELEPIYKFOVGLKINLOEEHVMALICLSPRPVODHARE 353
OY 364 QLOEOPATLKSYTECNROPAPHRELFELKIMAMTELRSINQHTO--RLKRIODIHP- 420
Db 354 ALDRLSELTQAVIYOLH--HPGRLLYAKMIQKIALDRSLRINEHRSKQYSLSPREHSMQ 411
OY 421 ATPMOELFG 430
Db 412 LTPVLEVSG 421

RESULT 14
ID 090U51 PRELIMINARY; PRT; 358 AA.
AC 090U51;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE NUCLEAR RECEPTOR.
GN CAR.
OS Rattus norvegicus (Rat.).
OC Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=WISTAR KYOTO, AND FISCHER;
RA Yoshinari K., Sueyoshi T., Moore R., Negishi M.;
RT "Sexually dimorphic nuclear translocation of receptor CAR and
RT induction of cyf2b1 gene by phenobarbital in rat livers."
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ133095; AAF22567.1; -
DR EMBL: AF133094; AAF22566.1; -
DR HSSP: P19793; 2ML.
DR INTERPRO: IPR000324; -
DR INTERPRO: IPR000536; -
DR INTERPRO: IPR001628; -
DR INTERPRO: IPR001723; -
DR INTERPRO: IPR001728; -
DR PFAM: PF00104; hormone_rec. 1.
DR PFAM: PF00105; zf-C4; 1.
DR PRINTS: PR00047; STROIDEFINGER.
DR PRINTS: PR00350; VITAMINDR.
DR PRINTS: PR00398; STRDHOMONER.
DR PRINTS: PR00546; THYROIDHORM.
DR PROSITE: PS00031; NUCLEAR_RECEPTOR; 1.
KW Receptor.
SQ SEQUENCE 358 AA; 40922 MW; 13691F49CAD8F1ED CRC64;

Query Match 31.8%; Score 729; DB 11; Length 358;
Best Local Similarity 40.3%; Pred. No. 1.4e-53;
Matches 160; Conservative 61; Mismatches 122; Indels 54; Gaps 6;

```

```

OY 273 ISLKGAPEFCOLRENTVENAETGMEGCRSLSYCLEDTAG--GFOQLLEPMKFKHYMLK 331
Db 201 ISLKGAPEVETLISLTTFCLOTFONFCPLCYKEDAVHGFQFELELIHFHFKTLK 260
OY 332 KLOHEEYVMOAISLSPDRPGVLOHRYVDLOEOPATLKSYTECNROPAPHRELFEL 391
Db 261 RLOEPEPYALMAMALFSPDRPGVTOREIIDLQEEVALIINNHMEQOSRLQSSFLYA 320
OY 392 KIMAMTELRSINQHTORLRIODIHPFATPLMOEL 428
Db 321 KIMGLAEALRSINSAYSIEHRIQGLSAM--MPLGEL 356

RESULT 15
ID 090UH7 PRELIMINARY; PRT; 445 AA.
AC 090UH7;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE NUCLEAR OXYSTEROL RECEPTOR LXR-ALPHA (LXR-ALPHA).
GN NR1H3 OR LXR-ALPHA.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129SVJ; TISSUE=LIVER;
RA Albert S., Steffensen K.R., Gustafsson J.A.;
RT "Cloning and characterisation of nuclear oxysterol receptor genes LXRa
RT and Lxrb from mouse."
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ132599; CAB51952.1; -
DR EMBL: AJ132600; CAB51952.1; JOINED.
DR EMBL: AJ132601; CAB51953.1; -
DR HSSP: P20393; 1A6Y.
DR MGD: MGI:1352462; Nr1h3.
DR INTERPRO: IPR000324; -
DR INTERPRO: IPR000536; -
DR INTERPRO: IPR000923; -
DR INTERPRO: IPR001628; -
DR INTERPRO: IPR001723; -
DR INTERPRO: IPR001728; -
DR INTERPRO: IPR003069; -
DR INTERPRO: IPR003078; -
DR PFAM: PF00104; hormone_rec. 1.
DR PFAM: PF00105; zf-C4; 1.
DR PRINTS: PR00047; STROIDEFINGER.
DR PRINTS: PR00350; VITAMINDR.
DR PRINTS: PR00398; STRDHOMONER.
DR PRINTS: PR00546; THYROIDHORM.
DR PRINTS: PR01283; ECDYSTEROIDR.
DR PRINTS: PR01282; RETINOICACIDR.
DR PROSITE: PS00031; NUCLEAR_RECEPTOR; 1.
DR PROSITE: PS00196; COPPER_BLUE; UNKNOWN_1.
KW Receptor.
SQ SEQUENCE 445 AA; 50476 MW; CC9ABDF38D935593 CRC64;

Query Match 22.0%; Score 504.5; DB 11; Length 445;
Best Local Similarity 30.4%; Pred. No. 1.6e-34;
Matches 130; Conservative 85; Mismatches 142; Indels 71; Gaps 13;

```

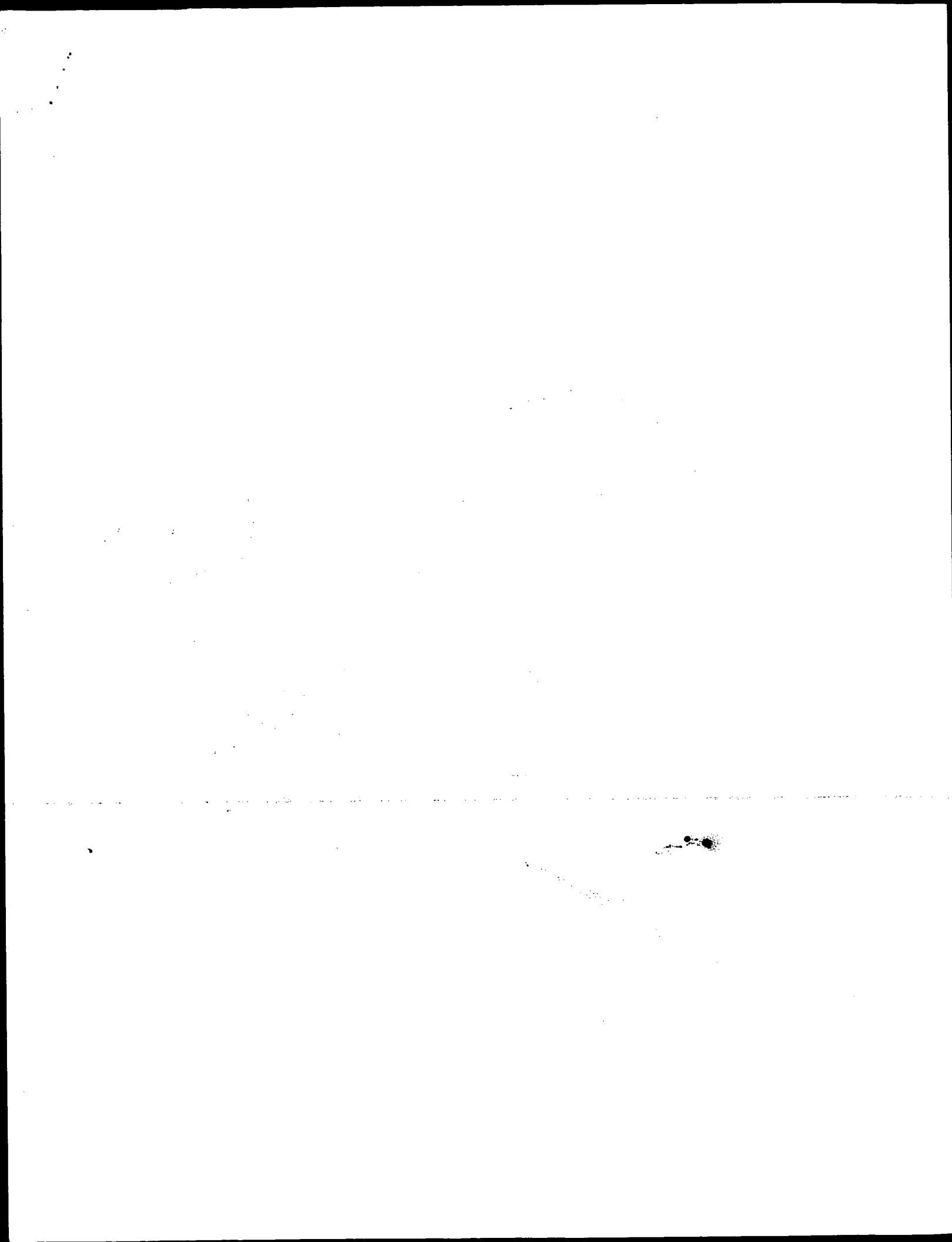


```

QY 132 ---RTGTOPLGVOGLTEORRMRELMDAOKTFDTTFSHFKNRPLPGVLSGCELPESL 188
Db 190 VSPRVSSPPQVLPQLSPQLEMIKELVAAOOCNRSFS----- 228
QY 189 QAPSRFAAKKSOVRKDCSLKVSLOLRGSDSVWNYKPPADSGKREIFSLPHMADMT 248
Db 229 ---DRLATYPR-PIAPD-----POSREARQORFA--HTELA 260
QY 249 YMEGIIISFAKVISYFRDLPIDQISLLKGAPELCLRENTVEN--AETGTECGRLSY 306
Db 261 VSVQEIYDFAKOLPGFQLSREDQIALKTSALIEVMLETSRRYNPGSESITF-LKDFSY 319
QY 307 CLEPTA-GFQOQLLEPMKLFHYMLKQLHEEYVLMQALISLSPDRPGVLOHRVVDOL 365
Db 320 NREDFAKAGQVEFINDIFEFSSRAMMELQINDAEFALLIAISIFSADRPVQDQLQVERL 379
QY 366 QEQFAITLKSIECNRPQAPAHREFLFLKIMAMTELRSINAOHQRL--LRIDIHFPATP 423
Db 380 QHTYVEALHAYVSINHPH--DRLMFPMLMKLVSLRTLSVHSEQVFAKRLQD--KKLPP 435
QY 424 LMOELFGI 431
Db 436 LTSEIMDV 443

```

Search completed: February 18, 2001, 13:05:04
 Job time: 6193 sec



GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 18, 2001, 12:12:43 ; Search time 63.99 Seconds

(without alignments)
231.913 Million cell updates/sec

Title: US-09-143-828-2

Perfect score: 2290

Sequence: 1 MEYRPRESNMHADVHCEDT.....ODIHPRATPLMQELFCITGS 434

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 268485 seqs, 34193795 residues

Total number of hits satisfying chosen parameters: 268485

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database:

1: /cgn2_2/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /cgn2_2/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /cgn2_2/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /cgn2_2/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /cgn2_2/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /cgn2_2/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /cgn2_2/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /cgn2_2/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /cgn2_2/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /cgn2_2/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /cgn2_2/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /cgn2_2/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /cgn2_2/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /cgn2_2/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /cgn2_2/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /cgn2_2/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /cgn2_2/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /cgn2_2/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /cgn2_2/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /cgn2_2/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /cgn2_2/gcgdata/geneseq/geneseq/AA2000.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	2290	100.0	437	20	Y09515 Human vitamin D re
2	2287	98.9	434	20	Y15931 A human intranucle
3	2287	98.9	457	20	Y16035 A human intranucle
4	2287	99.9	457	20	Y15932 A human intranucle
5	2287	99.9	466	20	Y25410 Human nmr7 partial
6	2287	99.9	473	20	Y25411 Human nmr7-1 prote
7	2287	99.9	473	20	Y15936 A human intranucle
8	2287	99.9	473	20	Y15933 A human intranucle
9	2287	99.9	473	20	Y09516 Human vitamin D re
10	2169	94.7	434	20	Y21799 Human steroid and
11	2165	94.5	414	20	Y42691 Human pregnane X r
12	1585	69.2	316	20	Y42689 H1s6-pregnane X re

13	979.5	42.8	386	17	R98521
14	801	35.0	423	20	W94623
15	798	34.8	423	19	W47509
16	789	34.5	427	19	W68156
17	789	34.5	427	20	Y09064
18	789	34.5	450	20	Y09036
19	789	34.5	477	20	Y09035
20	768	33.5	348	18	W32536
21	768	33.5	348	20	W93902
22	763	33.3	348	14	R41346
23	733.5	32.0	357	20	Y17872
24	725	31.7	358	20	W93903
25	656	28.6	356	19	W37261
26	656	28.6	356	20	W94622
27	632	27.6	367	14	R43656
28	498.5	21.8	445	21	Y32374
29	498	21.7	460	16	R74738
30	498	21.7	461	15	R32980
31	498	21.7	461	17	R97982
32	498	21.7	461	17	R98140
33	496	21.7	460	18	W25034
34	495	21.6	461	17	R96234
35	491.5	21.5	447	17	W03326
36	490	21.4	446	17	R99736
37	487	21.3	446	17	R94169
38	483	21.1	443	16	R74739
39	483	21.1	443	18	W25035
40	481.5	21.0	440	14	R33744
41	469.5	20.5	472	19	W40072
42	454	19.8	469	17	W03448
43	451.5	19.7	484	9	R99735
44	448	19.6	456	17	P80921
45	445	19.4	461	20	Y21631

ALIGNMENTS

RESULT	1
ID	Y09515
Y09515	standard; Protein: 437 AA.
XX	Y09515:
AC	XX
DT	16-JUL-1999 (first entry)
XX	XX
DE	Human vitamin D receptor related gamma protein.
XX	XX
KW	Human; vitamin D receptor related protein; VDR; obesity; diabetes; anorexia; rheumatoid arthritis; lipoprotein defect; hyperlipidaemia; hypercholesterolaemia; hyperlipoproteinaemia; osteoporosis; tumour; hyperproliferative skin disorder; hyperthyroidism.
KW	Homo sapiens.
OS	XX
PN	W09919354-A1.
XX	XX
PD	22-APR-1999.
XX	XX
PF	31-AUG-1998; 98WO-SE01548.
XX	XX
PR	31-MAR-1998; 98SF-0001148.
XX	XX
PR	14-OCT-1997; 97SE-0003745.
XX	XX
PA	(PHAA) PHARMACIA & UPJOHN AB.
XX	XX
PI	Berkenstam A, Dahlberg M;
XX	XX
DR	WPI; 1999-302508/25.
XX	XX
PT	New vitamin D receptor related (VDR) polypeptides, useful for treating obesity, diabetes, anorexia and rheumatoid arthritis

Xenopus orphan rec
Rat vitamin D rece
Rat vitamin D rece
Human vitamin D re
Human vitamin D re
Human vitamin D re
Human vitamin D re
Constitutively act
Human CAR receptor
Human CAR receptor
Mouse nuclear rece
Mouse CAR receptor
Rat vitamin D rece
Rat vitamin D rece
Mouse CNREB-1. Mu
Human ubiquitinous n
Human recombinant
Human steroid rece
NER receptor poten
Human foetal lung
LXR-alpha, orphan
Retinoid X recepto
OR-1 orphan recept
Rat ubiquitinous nuc
XR2. Homo sapiens
Human retinoid rec
Farnesoid-activate
Retinoid X recepto
Sequence encoded b
Ligand binding dom

XX Claim 19; Page 19-20; 35pp; English.
 CC The present sequence is a human vitamin D receptor related (VDR)
 CC polypeptide. Human VDR polypeptides and substances which affect VDR
 CC signal transduction, can be used for treating metabolic, proliferative
 CC or inflammatory conditions. They can be used in the manufacture of a
 CC medicament for treating the following conditions: obesity, diabetes,
 CC anorexia, lipoprotein defects, hyperlipidaemia, hypercholesterolaemia or
 CC hyperlipoproteinaemia and osteoporosis, rheumatoid arthritis, Dengue and
 CC malignant tumours, hyperproliferative skin disorders or hyperthyroidism.
 CC Nucleic acid vectors encoding for expression of a VDR polypeptide can
 CC be used for treating metabolic, proliferative or inflammatory conditions,
 CC by introducing them into a mammal. The introduced nucleic acid is then
 CC capable of transforming a cell in vitro and then polypeptide is expressed.
 CC A substance affecting VDR signal transduction, such as an agonist or
 CC antagonist can be used for the manufacture of a medicament for treating
 CC metabolic, proliferative or inflammatory condition.
 CC N.B. The specification specifically claims the VDR nucleic acid and
 CC polypeptide sequences given in figures 1, 4, 7 and 8, but no figures
 CC are given in the specification.

XX Sequence 437 AA;

Query Match 100.0%; Score 2290; DB 20; Length 437;
 Best Local Similarity 100.0%; Pred. No. 7.7e-219;
 Matches 434; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEVPRKESWNHADFVCEDESVPGRPSVNADEVGPGQICRYCGDKATGYHFNWTCG 60
 DB 1 MEVPRKESWNHADFVCEDESVPGRPSVNADEVGPGQICRYCGDKATGYHFNWTCG 60
 QY 61 CCGFRRRAKRNARLRCPRRKACETTRKTRQCCACRLRKLCSGKMKEMMSDEAYVE 120
 DB 61 CCGFRRRAKRNARLRCPRRKACETTRKTRQCCACRLRKLCSGKMKEMMSDEAYVE 120
 QY 121 RRALIKRKKSERGTQPLVGQGLTEEOQRMMIRELMDAOKTEPDTFSHKNRRLPGVLS 180
 DB 121 RRALIKRKKSERGTQPLVGQGLTEEOQRMMIRELMDAOKTEPDTFSHKNRRLPGVLS 180
 QY 181 GCELPESLOAPSRERAAKMSQVRKDLCSLKSLSQLRGEGSVWYKPPADSGGKEIFSL 240
 DB 181 GCELPESLOAPSRERAAKMSQVRKDLCSLKSLSQLRGEGSVWYKPPADSGGKEIFSL 240
 QY 241 PHMAAMSTYMERGIISFAVVISYFRDLPREDQISILKGAAPFLCQRRNTVNAETGWE 300
 DB 241 PHMAAMSTYMERGIISFAVVISYFRDLPREDQISILKGAAPFLCQRRNTVNAETGWE 300
 QY 301 CGRLSYCLEDTAGGFOQLLEPMLKFNHMLKRLQLEHEEYVLMQATSLFSPDRPGVLOHR 360
 DB 301 CGRLSYCLEDTAGGFOQLLEPMLKFNHMLKRLQLEHEEYVLMQATSLFSPDRPGVLOHR 360
 QY 361 VVDQLOEFAITLKSYIECNRPQAPARFLFLKIMAMTELRSINAOHTQRLRLRIDIHPE 420
 DB 361 VVDQLOEFAITLKSYIECNRPQAPARFLFLKIMAMTELRSINAOHTQRLRLRIDIHPE 420
 QY 421 ATPLMQELFGITGS 434
 DB 421 ATPLMQELFGITGS 434

RESULT 2

ID Y15931 standard; Protein; 434 AA.

AC Y15931;

DT 04-AUG-1999 (first entry)

DE A human intranuclear receptor protein.

XX Human; intranuclear receptor protein; drug development; diagnosis;

KW treatment.

XX OS Homo sapiens.

XX PN JP1127872-A.

XX PD 18-MAY-1999.

XX PF 07-AUG-1998; 98UP-0224172.

XX PR 11-AUG-1997; 97UP-0230335.

XX PA (NIRS) JAPAN TOBACCO, INC.

XX WP1: 1999-350330/30.

XX N-PSDB; X59966.

XX New intranuclear receptor protein - useful for drug development and

XX diagnosis and treatment of disease

XX Claim 1; Page 15-16; 38pp; Japanese.

XX The present sequence represents a human intranuclear receptor protein.

XX The nucleic acid sequence was isolated from a human adult cDNA

XX library using a swillfish AMO23 derived probe. The protein can

XX be used for the development of drugs and diagnosis and treatment

XX of various diseases.

XX Sequence 434 AA;

Query Match 99.9%; Score 2287; DB 20; Length 434;
 Best Local Similarity 99.8%; Pred. No. 1.5e-218;
 Matches 433; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEVPRKESWNHADFVCEDESVPGRPSVNADEVGPGQICRYCGDKATGYHFNWTCG 60
 DB 1 MEVPRKESWNHADFVCEDESVPGRPSVNADEVGPGQICRYCGDKATGYHFNWTCG 60
 QY 61 CCGFRRRAKRNARLRCPRRKACETTRKTRQCCACRLRKLCSGKMKEMMSDEAYVE 120
 DB 61 CCGFRRRAKRNARLRCPRRKACETTRKTRQCCACRLRKLCSGKMKEMMSDEAYVE 120
 QY 121 RRALIKRKKSERGTQPLVGQGLTEEOQRMMIRELMDAOKTEPDTFSHKNRRLPGVLS 180
 DB 121 RRALIKRKKSERGTQPLVGQGLTEEOQRMMIRELMDAOKTEPDTFSHKNRRLPGVLS 180
 QY 181 GCELPESLOAPSRERAAKMSQVRKDLCSLKSLSQLRGEGSVWYKPPADSGGKEIFSL 240
 DB 181 GCELPESLOAPSRERAAKMSQVRKDLCSLKSLSQLRGEGSVWYKPPADSGGKEIFSL 240
 QY 241 PHMAAMSTYMERGIISFAVVISYFRDLPREDQISILKGAAPFLCQRRNTVNAETGWE 300
 DB 241 PHMAAMSTYMERGIISFAVVISYFRDLPREDQISILKGAAPFLCQRRNTVNAETGWE 300
 QY 301 CGRLSYCLEDTAGGFOQLLEPMLKFNHMLKRLQLEHEEYVLMQATSLFSPDRPGVLOHR 360
 DB 301 CGRLSYCLEDTAGGFOQLLEPMLKFNHMLKRLQLEHEEYVLMQATSLFSPDRPGVLOHR 360
 QY 361 VVDQLOEFAITLKSYIECNRPQAPARFLFLKIMAMTELRSINAOHTQRLRLRIDIHPE 420
 DB 361 VVDQLOEFAITLKSYIECNRPQAPARFLFLKIMAMTELRSINAOHTQRLRLRIDIHPE 420
 QY 421 ATPLMQELFGITGS 434
 DB 421 ATPLMQELFGITGS 434

RESULT 3

ID Y16035 standard; Protein; 457 AA.

AC Y16035;

XX 04-AUG-1999 (first entry)

XX A human intranuclear receptor protein.

XX Human; intranuclear receptor protein; drug development; diagnosis;
treatment.

XX Homo sapiens.

XX JP1127872-A.

XX 18-MAY-1999.

XX 07-AUG-1998; 98JP-0224172.

XX 11-AUG-1997; 97JP-0230335.

XX (NISB) JAPAN TOBACCO INC.

XX WPI: 1999-350330/30.

XX N-PSDB; X59975.

XX New intranuclear receptor protein - useful for drug development and
diagnosis and treatment of disease

XX Disclosure: Page 35-37; 38pp; Japanese.

XX The present sequence represents a human intranuclear receptor protein.

XX The nucleic acid sequence was isolated from a human adult cDNA

XX library using a swillfish ANO23 derived probe. The protein can

XX be used for the development of drugs and diagnosis and treatment

XX of various diseases.

XX Sequence 457 AA;

XX Query Match 99.9%; Score 2287; DB 20; Length 457;
Best Local Similarity 99.8%; Pred. No. 1.6e-218;
Matches 433; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEVPRKESNMHADFYHCEDETSVPGRPSVNADEVGPGQICRVCGDKATGYHFNMTCEG 60

DB 24 LEVRPKESNMHADFYHCEDETSVPGRPSVNADEVGPGQICRVCGDKATGYHFNMTCEG 83

QY 61 CKGFRRAMKRNARLRCPPRKACETTRKTRQCCACRLKCLSEGMKEMTSDAIVEE 120

DB 84 CKGFRRAMKRNARLRCPPRKACETTRKTRQCCACRLKCLSEGMKEMTSDAIVEE 143

QY 121 RRALKRKRSERTGTOPGLVGGLTGEORMMIRELMAOKTFPTTSHKFNRLPGVLS 180

DB 144 RRALKRKRSERTGTOPGLVGGLTGEORMMIRELMAOKTFPTTSHKFNRLPGVLS 203

QY 181 GCELPESLQAPSRREBAKMSQVRKDCSLKVSILQRLGEDGSVWNYKPPADSGKEIFSL 240

DB 204 GCELPESLQAPSRREBAKMSQVRKDCSLKVSILQRLGEDGSVWNYKPPADSGKEIFSL 263

QY 241 PHMADSTYMFKGIIISFAKVISYFRDLPIEDQISLKGAAFLCOLRNTVFAETGWE 300

DB 264 PHMADSTYMFKGIIISFAKVISYFRDLPIEDQISLKGAAFLCOLRNTVFAETGWE 323

QY 301 CGRLSTCLEDLAGGFOQLLEPMLKFHYMLKRLQLEHEEYVLMQATSLSPDRPGVLOHR 360

DB 324 CGRLSTCLEDLAGGFOQLLEPMLKFHYMLKRLQLEHEEYVLMQATSLSPDRPGVLOHR 383

QY 361 VVDOLQOPAITLKSYTECNRPQAPARFLFLKIMAMLTSLRSINQHTORLRIODIHPF 420

DB 384 VVDOLQOPAITLKSYTECNRPQAPARFLFLKIMAMLTSLRSINQHTORLRIODIHPF 443

QY 421 ATPLMOELFGITGS 434

DB 444 ATPLMOELFGITGS 457

RESULT 4

XX Y15932

XX Y15932 standard; Protein: 457 AA.

XX Y15932;

XX 04-AUG-1999 (first entry)

XX A human intranuclear receptor protein.

XX Human; intranuclear receptor protein; drug development; diagnosis;
treatment.

XX Homo sapiens.

XX JP1127872-A.

XX 18-MAY-1999.

XX 07-AUG-1998; 98JP-0224172.

XX 11-AUG-1997; 97JP-0230335.

XX (NISB) JAPAN TOBACCO INC.

XX WPI: 1999-350330/30.

XX N-PSDB; X59967.

XX New intranuclear receptor protein - useful for drug development and
diagnosis and treatment of disease

XX Claim 2; Page 16-17; 38pp; Japanese.

XX The present sequence represents a human intranuclear receptor protein.

XX The nucleic acid sequence was isolated from a human adult cDNA

XX library using a swillfish ANO23 derived probe. The protein can

XX be used for the development of drugs and diagnosis and treatment

XX of various diseases.

XX Sequence 457 AA;

XX Query Match 99.9%; Score 2287; DB 20; Length 457;
Best Local Similarity 99.8%; Pred. No. 1.6e-218;
Matches 433; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEVPRKESNMHADFYHCEDETSVPGRPSVNADEVGPGQICRVCGDKATGYHFNMTCEG 60

DB 24 LEVRPKESNMHADFYHCEDETSVPGRPSVNADEVGPGQICRVCGDKATGYHFNMTCEG 83

QY 61 CKGFRRAMKRNARLRCPPRKACETTRKTRQCCACRLKCLSEGMKEMTSDAIVEE 120

DB 84 CKGFRRAMKRNARLRCPPRKACETTRKTRQCCACRLKCLSEGMKEMTSDAIVEE 143

QY 121 RRALKRKRSERTGTOPGLVGGLTGEORMMIRELMAOKTFPTTSHKFNRLPGVLS 180

DB 144 RRALKRKRSERTGTOPGLVGGLTGEORMMIRELMAOKTFPTTSHKFNRLPGVLS 203

QY 181 GCELPESLQAPSRREBAKMSQVRKDCSLKVSILQRLGEDGSVWNYKPPADSGKEIFSL 240

DB 204 GCELPESLQAPSRREBAKMSQVRKDCSLKVSILQRLGEDGSVWNYKPPADSGKEIFSL 263

QY 241 PHMADSTYMFKGIIISFAKVISYFRDLPIEDQISLKGAAFLCOLRNTVFAETGWE 300

DB 264 PHMADSTYMFKGIIISFAKVISYFRDLPIEDQISLKGAAFLCOLRNTVFAETGWE 323

QY 301 CGRLSTCLEDLAGGFOQLLEPMLKFHYMLKRLQLEHEEYVLMQATSLSPDRPGVLOHR 360

DB 324 CGRLSTCLEDLAGGFOQLLEPMLKFHYMLKRLQLEHEEYVLMQATSLSPDRPGVLOHR 383

QY 361 VVDOLQOPAITLKSYTECNRPQAPARFLFLKIMAMLTSLRSINQHTORLRIODIHPF 420

Db 384 vvdqieqfaatlksylecnpqphrflfkimamltelrsinaqhtqrlrlrldhpf 443
 Oy 421 ATPLMOELFGITGS 434
 Db 444 atplmqelfgltgs 457

RESULT 5
 Y25410
 ID Y25410 standard; Protein; 466 AA.

AC Y25410;
 DT 06-SEP-1999 (first entry)
 DE Human nNR7 partial protein.

XX nNR7; nNR7-1; nuclear trans-acting receptor protein; human; regulator;
 KW identification; downstream target gene; cell proliferation;
 cell development.

OS Homo sapiens.

PN MO9931129-A1.

PD 24-JUN-1999.

PE 11-DEC-1998; 98WO-US26364.

PR 14-OCT-1998; 98US-0104251.

PR 12-DEC-1997; 97US-0069401.

PA (MERI) MERCK & CO INC.

PI Chen F;

DR WPI; 1999-405024/34.

DR N-PSDB; X78807.

PT DNA encoding human nuclear receptors nNR7 and nNR7-1

PS Claim 1; Fig 3; 80pp; English.

XX This invention describes the isolation of the novel human nuclear
 CC receptors nNR7 or nNR7-1. The nNR7 and nNR7-1 proteins are useful in
 CC the identification of downstream target genes and ligands regulating its
 CC activity. The nuclear receptor is involved in the regulation of in vivo
 CC cell proliferation and/or cell development. The nNR7 and nNR7-1
 CC polynucleotides, expression vectors and host cells are useful for the
 CC recombinant production of the protein.

XX Sequence 466 AA;

Query Match 99.9%; Score 2287; DB 20; Length 466;
 Best Local Similarity 99.8%; Pred. No. 1.7e-218;
 Matches 433; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 33 levprkeswnhadvfhcedtesvpgkpsvnadeevgpgqicrvgdktatgyhfnvmcteg 92
 Oy 1 MEVPRKESWNHADVHCEDESVPKESVNADEYVGPGQICRVGDKATGYHFNVMCTEG 60
 : |||||
 Db 61 CKGFERRAMKRNARLRCPFRKACEITRKTRROCAQRLRKCSGKKKEMINSDEAVEE 120
 : |||||
 Db 93 ckgffrramkrnarlrpfrkaceitrtkrtrrocaqrlrkcslesgmkkeminsdeavee 152
 : |||||
 Oy 121 RRALIKRKRKSERGTQPLGVGLTEQRMIRRLMDAQMFTFTTSHFNFRLLPGVLSS 180
 : |||||
 Db 153 rralikrkrkeritgtplgyglteeqrmrlrelmdagmkttdttfhnfrllpyvlss 212
 : |||||
 Oy 181 GCELPESIQAPSRREAAKMSQVRDCLSLKYSLOLRGEDSVWNYKRPADSGKEIFSL 240
 : |||||
 Db 213 gceipesiqapsreearakmsqvrkdclslkyslqlrgedsvwnykpadsgekeifsl 272

Oy 241 PHMADSTYMEKGIISFAKVISYFRDLPIEDQISLKGAFELCOLRNTVFNATGTWE 300
 : |||||
 Db 273 pnmadstymefkylsfakvisyfrdlpiedqisllkgaefelcolrntvfnatgtwe 332
 : |||||
 Oy 301 CGRLSYCLEDETFAGGFQOLLLEPMLKFRHYMKRIQHEEYVLMQAIISPSDRPGVLOHR 360
 : |||||
 Db 333 cgrlsycaledtfaggfqqllepmlkfhymkkrkqlheeyvlmqaisfspdrgvlgqr 392
 : |||||
 Oy 361 VVDQLOEFAFIRLTKSTIECNRPQPARFRFLKIMAMLTFLRSINQHTORLLRIODHPF 420
 : |||||
 Db 393 vvdqieqfaatlksylecnpqphrflfkimamltelrsinaqhtqrlrlrldhpf 452
 : |||||
 Oy 421 ATPLMOELFGITGS 434
 Db 453 atplmqelfgltgs 466

RESULT 6
 Y25411
 ID Y25411 standard; Protein; 473 AA.

AC Y25411;

DT 06-SEP-1999 (first entry)

DE Human nNR7-1 protein.

XX nNR7; nNR7-1; nuclear trans-acting receptor protein; human; regulator;
 KW identification; downstream target gene; cell proliferation;
 cell development.

OS Homo sapiens.

PN WO9931129-A1.

PD 24-JUN-1999.

PE 11-DEC-1998; 98WO-US26364.

PR 14-OCT-1998; 98US-0104251.

PR 12-DEC-1997; 97US-0069401.

PA (MERI) MERCK & CO INC.

PI Chen F;

DR WPI; 1999-405024/34.

DR N-PSDB; X78808.

PT DNA encoding human nuclear receptors nNR7 and nNR7-1

PS Claim 20; Fig 6; 80pp; English.

XX This invention describes the isolation of the novel human nuclear
 CC receptors nNR7 or nNR7-1. The nNR7 and nNR7-1 proteins are useful in
 CC the identification of downstream target genes and ligands regulating its
 CC activity. The nuclear receptor is involved in the regulation of in vivo
 CC cell proliferation and/or cell development. The nNR7 and nNR7-1
 CC polynucleotides, expression vectors and host cells are useful for the
 CC recombinant production of the protein.

XX Sequence 473 AA;

Query Match 99.9%; Score 2287; DB 20; Length 473;
 Best Local Similarity 99.8%; Pred. No. 1.7e-218;
 Matches 433; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 40 levprkeswnhadvfhcedtesvpgkpsvnadeevgpgqicrvgdktatgyhfnvmcteg 99
 Oy 1 MEVPRKESWNHADVHCEDESVPKESVNADEYVGPGQICRVGDKATGYHFNVMCTEG 60
 : |||||

QY 61 CKGFFRRAMKRNARLRCPFRKGCETTRKTRQOACRLKCLSGMKKEMIMSDAEVE 120
 DB 100 CCGFFRRAMKRNARLRCPFRKGCETTRKTRQOACRLKCLSGMKKEMIMSDAEVE 159
 QY 121 RRALIRKRSERSTGTPLOGLTEGQRMIRLMDAQKTFDTTSHKRNFLPGVLS 180
 DB 160 RRALIRKRSERSTGTPLOGLTEGQRMIRLMDAQKTFDTTSHKRNFLPGVLS 219
 QY 181 GCELPESLQAPSRREAAKMSOVRKDCSLKVSLOLGEOSVWNYKPPADSGKEIFSL 240
 DB 220 GCELPESLQAPSRREAAKMSOVRKDCSLKVSLOLGEOSVWNYKPPADSGKEIFSL 279
 QY 241 PHADMSTYFKGISFAVYISFRDLPEDQISLKGAFELCOLRFNTVFNAETGWE 300
 DB 280 PHADMSTYFKGISFAVYISFRDLPEDQISLKGAFELCOLRFNTVFNAETGWE 339
 QY 301 CGRLSYCLEDTAGGFOQLLEPMLKHYMLKRLQALHEEYVLMQALSLSPDRPGVLQHR 360
 DB 340 CGRLSYCLEDTAGGFOQLLEPMLKHYMLKRLQALHEEYVLMQALSLSPDRPGVLQHR 399
 QY 361 VVDQLOEOPAITLKSYTECNRPQAPARHFLFLKIMMTELRSINQHTORLRIODIHPF 420
 DB 400 VVDQLOEOPAITLKSYTECNRPQAPARHFLFLKIMMTELRSINQHTORLRIODIHPF 459
 QY 421 ATPLMQELFGITGS 434
 DB 460 ATPLMQELFGITGS 473

RESULT 7

ID Y15936 standard; Protein: 473 AA.

XX Y15936;

DT 04-AUG-1999 (first entry)

DE A human intranuclear receptor protein.

KW Human; intranuclear receptor protein; drug development; diagnosis;

OS Homo sapiens.

PN JP1127872-A.

PD 18-MAY-1999.

PF 07-AUG-1998; 98JP-0224172.

PR 11-AUG-1997; 97JP-0230335.

PA (NISB) JAPAN TOBACCO INC.

DR WPI: 1999-350330/30.

DR N-PSDB; X59974.

PT New intranuclear receptor protein - useful for drug development and

PS diagnosis and treatment of disease

XX Disclosure: Page 32-35; 38pp; Japanese.

CC The present sequence represents a human intranuclear receptor protein.

CC The nucleic acid sequence was isolated from a human adult cDNA

CC library using a swedish ANO23 derived probe. The protein can

CC be used for the development of drugs and diagnosis and treatment

CC of various diseases.

SO Sequence 473 AA;

Query Match 99.9%; Score 2287; DB 20; Length 473;
 Best Local Similarity 99.8%; Pred. NO. 1.7e-218;

Matches 433; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MEVPRKSNMADPVEDPESVGRKSVNADVEVGPOICRVCGSKAGYHFNWTCG 60
 DB 40 Ievprksnmadpvedpessvgrksvvnadevvgpqcrcvgsdkaqyhnwmcg 99
 QY 61 CKGFFRRAMKRNARLRCPFRKGCETTRKTRQOACRLKCLSGMKKEMIMSDAEVE 120
 DB 100 CCGFFRRAMKRNARLRCPFRKGCETTRKTRQOACRLKCLSGMKKEMIMSDAEVE 159
 QY 121 RRALIRKRSERSTGTPLOGLTEGQRMIRLMDAQKTFDTTSHKRNFLPGVLS 180
 DB 160 RRALIRKRSERSTGTPLOGLTEGQRMIRLMDAQKTFDTTSHKRNFLPGVLS 219
 QY 181 GCELPESLQAPSRREAAKMSOVRKDCSLKVSLOLGEOSVWNYKPPADSGKEIFSL 240
 DB 220 GCELPESLQAPSRREAAKMSOVRKDCSLKVSLOLGEOSVWNYKPPADSGKEIFSL 279
 QY 241 PHADMSTYFKGISFAVYISFRDLPEDQISLKGAFELCOLRFNTVFNAETGWE 300
 DB 280 PHADMSTYFKGISFAVYISFRDLPEDQISLKGAFELCOLRFNTVFNAETGWE 339
 QY 301 CGRLSYCLEDTAGGFOQLLEPMLKHYMLKRLQALHEEYVLMQALSLSPDRPGVLQHR 360
 DB 340 CGRLSYCLEDTAGGFOQLLEPMLKHYMLKRLQALHEEYVLMQALSLSPDRPGVLQHR 399
 QY 361 VVDQLOEOPAITLKSYTECNRPQAPARHFLFLKIMMTELRSINQHTORLRIODIHPF 420
 DB 400 VVDQLOEOPAITLKSYTECNRPQAPARHFLFLKIMMTELRSINQHTORLRIODIHPF 459
 QY 421 ATPLMQELFGITGS 434
 DB 460 ATPLMQELFGITGS 473

RESULT 8

ID Y15933 standard; Protein: 473 AA.

XX Y15933;

DT 04-AUG-1999 (first entry)

DE A human intranuclear receptor protein.

KW Human; intranuclear receptor protein; drug development; diagnosis;

OS Homo sapiens.

PN JP1127872-A.

PD 18-MAY-1999.

PF 07-AUG-1998; 98JP-0224172.

PR 11-AUG-1997; 97JP-0230335.

PA (NISB) JAPAN TOBACCO INC.

DR WPI: 1999-350330/30.

DR N-PSDB; X59968.

PT New intranuclear receptor protein - useful for drug development and

PS diagnosis and treatment of disease

XX Claim 2; Page 17-19; 38pp; Japanese.

CC The present sequence represents a human intranuclear receptor protein.

CC The nucleic acid sequence was isolated from a human adult cDNA

CC library using a swedish ANO23 derived probe. The protein can

CC be used for the development of drugs and diagnosis and treatment

CC of various diseases.

XX Sequence 473 AA;

Query Match Best Local Similarity 99.9%; Score 2287; DB 20; Length 473;
Matches 433; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEVPRKESNMHADVHCEDTESVPGKPSVNADEEVGPOICRVCGDKATGYHFNVMTCG 60
:|||||
DB 40 LEVPRKESNMHADVHCEDTESVPGKPSVNADEEVGPOICRVCGDKATGYHFNVMTCG 99
QY 61 CKGFRRAMRNARLPCFPKRGACETIRKTRROCCARLRCLESQMKKEMIMSDAEVE 120
:|||||
DB 100 CKGFRRAMRNARLPCFPKRGACETIRKTRROCCARLRCLESQMKKEMIMSDAEVE 159
QY 121 RRALIRKKKSEPTGTOPLGVOGLTEQORMMIRELMDAOMKFTDTTFSHFNRLPGVLS 180
:|||||
DB 160 RRALIRKKKSEPTGTOPLGVOGLTEQORMMIRELMDAOMKFTDTTFSHFNRLPGVLS 219
QY 181 GCELPESLQAPSRREAAKMSQVRKDLCSLKVSLQLRGDSVWNYKPPADSGKEIFSL 240
:|||||
DB 220 GCELPESLQAPSRREAAKMSQVRKDLCSLKVSLQLRGDSVWNYKPPADSGKEIFSL 279
QY 241 PHMADSTYMFKGIISFAKVISYFEDLPIDQISLKGAEELCOLRENTVFNAETGWE 300
:|||||
DB 280 PHMADSTYMFKGIISFAKVISYFEDLPIDQISLKGAEELCOLRENTVFNAETGWE 339
QY 301 CGRLSTCLEDTAGFQOULLPEMLKFHYMLKQLHEEYVLMQALISFSDRPGVLOHR 360
:|||||
DB 340 CGRLSTCLEDTAGFQOULLPEMLKFHYMLKQLHEEYVLMQALISFSDRPGVLOHR 399
QY 361 VTDOLQEOFAITLKSYIECNRPQPAHRLFLKIMAMTELRISINAQHTQLRIODIHPE 420
:|||||
DB 400 VTDOLQEOFAITLKSYIECNRPQPAHRLFLKIMAMTELRISINAQHTQLRIODIHPE 459
QY 421 ATPLMQELFGITGS 434
:|||||
DB 460 ATPLMQELFGITGS 473

RESULT 9
ID Y09516 standard; Protein; 473 AA.

AC Y09516;
DT 16-JUL-1999 (first entry)
DE Human vitamin D receptor related gamma 2 protein.
KW Human; vitamin D receptor related protein; VDR; obesity; diabetes;
KW anorexia; rheumatoid arthritis; lipoprotein defect; hyperlipidaemia;
KW hypercholesterolaemia; hyperlipoproteinaemia; osteoporosis; tumour;
KW hyperproliferative skin disorder; hyperthyroidism.
OS Homo sapiens.
PN MO9919354-A1.
PD 22-APR-1999.
PE 31-AUG-1998; 98MO-SE01548.
PR 31-MAR-1998; 98SE-0001148.
PR 14-OCT-1997; 97SE-0003745.
PA (PMAA) PHARMACIA & UPJOHN AB.
PI Berkenstam A, Dahlberg M;
WP1; 1999-302508/25.
DR N-PSDB; X56243.

XX New vitamin D receptor related (VDR) polypeptides, useful for
PT treating obesity, diabetes, anorexia and rheumatoid arthritis
XX Claim 19; Page 22-24; 35pp; English.

CC The present sequence is a human vitamin D receptor related (VDR)
CC polypeptide. Human VDR polypeptides and substances which affect VDR
CC signal transduction, can be used for treating metabolic, proliferative
CC or inflammatory conditions. They can be used in the manufacture of a
CC medicament for treating the following conditions: obesity, diabetes,
CC anorexia, lipoprotein defects, hyperlipidaemia, hypercholesterolaemia or
CC hyperlipoproteinaemia and osteoporosis, rheumatoid arthritis, benign and
CC malign tumours, hyperproliferative skin disorders or hyperthyroidism.
CC Nucleic acid vectors encoding for expression of a VDR polypeptide can
CC be used for treating metabolic, proliferative or inflammatory conditions,
CC by introducing them into a mammal. The introduced nucleic acid is then
CC capable of transforming a cell in vivo and then polypeptide is expressed.
CC A substance affecting VDR signal transduction, such as an agonist or
CC antagonist can be used for the manufacture of a medicament for treating
CC metabolic, proliferative or inflammatory condition.
CC N.B. The specification specifically claims the VDR nucleic acid and
CC polypeptide sequences given in figures 1, 4, 7 and 8, but no figures
CC are given in the specification.

SO Sequence 473 AA;

Query Match Best Local Similarity 99.9%; Score 2287; DB 20; Length 473;
Matches 433; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEVPRKESNMHADVHCEDTESVPGKPSVNADEEVGPOICRVCGDKATGYHFNVMTCG 60
:|||||
DB 40 LEVPRKESNMHADVHCEDTESVPGKPSVNADEEVGPOICRVCGDKATGYHFNVMTCG 99
QY 61 CKGFRRAMRNARLPCFPKRGACETIRKTRROCCARLRCLESQMKKEMIMSDAEVE 120
:|||||
DB 100 CKGFRRAMRNARLPCFPKRGACETIRKTRROCCARLRCLESQMKKEMIMSDAEVE 159
QY 121 RRALIRKKKSEPTGTOPLGVOGLTEQORMMIRELMDAOMKFTDTTFSHFNRLPGVLS 180
:|||||
DB 160 RRALIRKKKSEPTGTOPLGVOGLTEQORMMIRELMDAOMKFTDTTFSHFNRLPGVLS 219
QY 181 GCELPESLQAPSRREAAKMSQVRKDLCSLKVSLQLRGDSVWNYKPPADSGKEIFSL 240
:|||||
DB 220 GCELPESLQAPSRREAAKMSQVRKDLCSLKVSLQLRGDSVWNYKPPADSGKEIFSL 279
QY 241 PHMADSTYMFKGIISFAKVISYFEDLPIDQISLKGAEELCOLRENTVFNAETGWE 300
:|||||
DB 280 PHMADSTYMFKGIISFAKVISYFEDLPIDQISLKGAEELCOLRENTVFNAETGWE 339
QY 301 CGRLSTCLEDTAGFQOULLPEMLKFHYMLKQLHEEYVLMQALISFSDRPGVLOHR 360
:|||||
DB 340 CGRLSTCLEDTAGFQOULLPEMLKFHYMLKQLHEEYVLMQALISFSDRPGVLOHR 399
QY 361 VTDOLQEOFAITLKSYIECNRPQPAHRLFLKIMAMTELRISINAQHTQLRIODIHPE 420
:|||||
DB 400 VTDOLQEOFAITLKSYIECNRPQPAHRLFLKIMAMTELRISINAQHTQLRIODIHPE 459
QY 421 ATPLMQELFGITGS 434
:|||||
DB 460 ATPLMQELFGITGS 473

RESULT 10
ID Y21799 standard; Protein; 434 AA.

AC Y21799;
DT 14-SEP-1999 (first entry)
XX

DE Human steroid and xenobiotic receptor (SXR).
 XX
 KW Nuclear receptor; SXR; steroid and xenobiotic receptor; RXR; human;
 KW retinoid X receptor; P450 gene; steroid hormone; steroid metabolism;
 KW phytoestrogen; calcium-channel blocker; steroid toxicity; tuberculosis;
 KW breast cancer; osteoporosis; Cushing syndrome; virilism; hirsutism;
 KW polycystic ovarian disease; cancer; colorectal; prostatic.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 227
 FT /label= unknown
 FT /note= "encoded by ACN"
 XX
 XX W09935246-A1.
 XX
 PD 15-JUL-1999.
 XX
 XX 08-JAN-1999; 99WO-US00490.
 XX
 XX 09-JAN-1998; 98US-0005286.
 XX
 XX (SALK) SALK INST BIOLOGICAL STUDIES.
 XX
 PI Blumberg B, Evans RM;
 XX
 DR WPI: 1999-419349/35.
 DR N-PSDB: X89080.
 XX
 PT New steroid and xenobiotic receptor, used to identify modulators for
 PT controlling metabolism of steroids and xenobiotics, e.g. reducing
 PT their toxicity
 XX
 PS Claim 4; Fig 1A; 83pp; English.
 XX
 XX The invention relates to a novel nuclear receptor polypeptide, designated
 CC SXR (steroid and xenobiotic receptor). SXR (1) forms a heterodimer with
 CC retinoid X receptor (RXR), (11) binds to a direct or inverted repeat
 CC response element motif based on the half-site AGTCA, (111) activates
 CC transcription through response elements present in steroid-inducible P450
 CC genes, in response to a wide variety of natural and synthetic steroid
 CC hormones and (1v) is prominently expressed in liver and intestine. SXR
 CC regulates expression of catabolic enzymes, in response to many different
 CC steroids, and thus affects metabolism. SXR is a broad specificity, low-
 CC affinity receptor for reducing excessive levels of steroids in the
 CC circulation. (Ant)agonists of SXR are used to regulate metabolism of
 CC steroids particularly phytoestrogens or calcium-channel blockers, to
 CC reduce steroid toxicity in subjects being treated with steroids, e.g. in
 CC cases of tuberculosis (treated with rifampin and related compounds),
 CC breast cancer (treated with tamoxifen, raloxifen etc.) or osteoporosis
 CC (treated with vitamin K), or to slow metabolism of therapeutic steroids.
 CC Also, modulating endogenous SXR is used to treat disease, particularly
 CC an agonist is used where endogenous steroid levels are excessive (e.g.
 CC Cushing syndrome; virilism and hirsutism in women; polycystic ovarian
 CC disease; 11 beta-, 17- or 21-hydroxylase deficiency; 3 beta-hydroxysteroid
 CC dehydrogenase deficiency, or breast, colorectal or prostatic cancer).
 CC While antagonists are used where endogenous steroid levels are too low.
 CC Cells that express SXR are used to identify compounds likely to be
 CC involved in undesirable drug interactions. Antibodies specific for SXR
 CC are used in immunohistochemical testing for studying distribution/
 CC expression density of SXR, also for diagnosis and therapeutically as
 CC antagonist. The present sequence represents SXR polypeptide.
 CC
 XX
 SO Sequence 434 AA:

Query Match 94.7%; Score 2169; DB 20; Length 434;
 Best Local Similarity 95.6%; Pred No. 7.7e-207;
 Matches 415; Conservative 3; Mismatches 16; Indels 0; Gaps 0;

1 MEVRPESNHNADPVHCDETSVPGKPSVNADEVGPOICRVCGDKATGYHFNMTCEG 60
 :|||||

DB 1 levypkeswnhadvhcedteavpkipsvnadeevygpqicrvcgdkatgyhfnmtceg 60
 OY 61 CKGFPRAMKRNRLCPFRKACETTRKTRRCCOACRLKCLSGMKKEMIMSDAIVE 120
 DB 61 CKGFPRAMKRNRLCPFRKACETTRKTRRCCOACRLKCLSGMKKEMIMSDAIVE 120
 OY 121 RRLIRKKKSERTGTPLGVGGLTEEDORMMIRELMDAOKRTDPTFSHRKRNRLPGVLSS 180
 DB 121 RRLIRKKKSERTGTPLGVGGLTEEDORMMIRELMDAOKRTDPTFSHRKRNRLPGVLSS 180
 OY 181 GCCLPESLAPSRERAAKSNQVAKDLCSIKVSIQLRGEGSVWNNKPPADSGKETFSLL 240
 DB 181 GCCLPESLAPSRERAAKSNQVAKDLCSIKVSIQLRGEGSVWNNKPPADSGKETFSLL 240
 OY 241 PHMADSTVFKGIISFARVSIYFRDLPIEDOSILKGAFAELCOLRFTVNAEGTWE 300
 DB 241 PHMADSTVFKGIISFARVSIYFRDLPIEDOSILKGAFAELCOLRFTVNAEGTWE 300
 OY 301 CGRLSYCLEDTAGCFQOOLLEPMLEFMYMLKRLQLHEEYVLMQALSLFSPDRPGVLQHR 360
 DB 301 CGRLSYCLEDTAGCFQOOLLEPMLEFMYMLKRLQLHEEYVLMQALSLFSPDRPGVLQHR 360
 OY 361 VVDQLQEQFAITTKSTIECKRFPQARHFLFKIMAMLTLSRSINAOHRLRIQIDHFF 420
 DB 361 VVDQLQEQFAITTKSTIECKRFPQARHFLFKIMAMLTLSRSINAOHRLRIQIDHFF 420
 OY 421 ATPLMDELFGITGS 434
 DB 421 ATPLMDELFGITGS 434
 DB 421 ATPLMDELFGITGS 434

RESULT 11
 ID Y42691
 AC Y42691;
 DT 17-JAN-2000 (first entry)
 DE Human pregnane X receptor (hPXR).
 XX
 KW Human; nuclear receptor; pregnane X receptor; PXR; CYP; CYP3A4;
 KW cytochrome P-450 mono-oxygenase; drug interaction; hPXR.
 XX
 OS Homo sapiens.
 XX
 PN W0948915-A1.
 PD 30-SEP-1999.
 PF 26-MAR-1999; 99WO-US06737.
 XX
 XX 27-MAR-1998; 98US-0079593.
 PR (GLAX) GLAXO GROUP LTD.
 PA
 PI Kilewer SA, Willson TM;
 XX
 DR WPI: 1999-601202/51.
 DR N-PSDB: Z07997.
 PT New human pregnane X receptor, used to identify specific modulators and
 PT agents that induce expression of cytochrome P-450 mono-oxygenase
 XX
 PS Claim 4; Fig 1A-D; 69pp; English.
 XX
 XX The invention provides an isolated human nuclear receptor (designated
 CC pregnane X receptor, PXR) that binds to a cytochrome P-450 mono-oxygenase
 CC (CYP) promoter. The hPXR is used to identify its specific modulators,
 CC and compounds that induce CYP3A4 expression (i.e. to identify drug
 CC interactions, since CYP3A4 is involved in many biotransformations of
 CC drugs). The modulators are potentially useful for: associating particular
 CC diseases and conditions with PXR and for treating such conditions.

CC Antibodies raised against hPXR can be used for determination and
CC purification of hPXR. The present sequence represents the hPXR.

XX
S0 Sequence 414 AA;

Query Match 94.5%; Score 2165; DB 20; Length 414;
Best Local Similarity 95.2%; Pred. No. 1.8e-206;
Matches 413; Conservative 1; Mismatches 0; Indels 20; Gaps 1;

QY 1 MEVPRKESNNADVHCEDTESVPEKPSVNADEYGVQICRVCGRKATGTHFNVMTCEG 60
DB :|||||
DB 1 LEVRPKESNNADVHCEDTESVPEKPSVNADEYGVQICRVCGRKATGTHFNVMTCEG 60
QY 61 CKGFPRBAMKRNARLRCFPRGACETTRTRCOACLRKCLSGMKEMKMSDEAVEE 120
DB :|||||
DB 61 CKGFPRBAMKRNARLRCFPRGACETTRTRCOACLRKCLSGMKEMKMSDEAVEE 120
QY 121 RRALIRRRKSEKSTGTPPLGVQGLTEEQRMIRLMDAQMKTFTTFSHKFNRLPGVLS 180
DB :|||||
DB 121 RRALIRRRKSEKSTGTPPLGVQGLTEEQRMIRLMDAQMKTFTTFSHKFNRLPGVLS 180
QY 181 GCELPESLQAPSRERAAWQVRKDCSLKVSLSLRGDSVWNTKPPADSGKEIFSL 240
DB :|||||
DB 181 GCELPESLQAPSRERAAWQVRKDCSLKVSLSLRGDSVWNTKPPADSGKEIFSL 240
QY 241 PHMADSTYMFKGIISFPAKVISYFRDLPIEDOISLKGAEFLCOLRNTYFNATGTCWE 300
DB :|||||
DB 241 PHMADSTYMFKGIISFPAKVISYFRDLPIEDOISLKGAEFLCOLRNTYFNATGTCWE 300
QY 301 CGRLSYCEDTAGFQQLLEBMLKFHYMLKKQLDHEEYVLMQALISFSPDRGVLOHR 360
DB :|||||
DB 301 CGRLSYCEDTAGFQQLLEBMLKFHYMLKKQLDHEEYVLMQALISFSPDRGVLOHR 360
QY 361 VVDQLQBPATLTKYIECNRPQPAHRLFLKIMAMLTSLNSINAGHTORLLRIQDIHPF 420
DB :|||||
DB 361 VVDQLQBPATLTKYIECNRPQPAHRLFLKIMAMLTSLNSINAGHTORLLRIQDIHPF 420
QY 421 ATPLMQELFGITGS 434
DB :|||||
DB 421 ATPLMQELFGITGS 434
QY 401 acplmqelfgitgs 414
DB :|||||
DB 401 acplmqelfgitgs 414

RESULT 12

Y42689
ID Y42689 standard; Protein; 316 AA.

AC Y42689;

DT 17-JAN-2000 (first entry)

DE H1s6-pregnane X receptor (PXR) sequence.

XX Human; nuclear receptor; pregnane X receptor; PXR; CYP; CYP3A4;
XX cytochrome P-450 mono-oxygenase; drug interaction; hPXR.

OS Synthetic.
XX Homo sapiens.

PN W09948915-A1.

PD 30-SEP-1999.

PE 26-MAR-1999; 99WO-US06737.

PR 27-MAR-1998; 98US-0079593.

PA (GLAXO) GLAXO GROUP LTD.

PI Kilewer SA, Willson TM;

DR WPI; 1999-601202/51.

XX

PT New human pregnane X receptor, used to identify specific modulators and
PT agents that induce expression of cytochrome P-450 mono-oxygenase

XX
PS Example 6; Page 37; 69pp; English.

CC The invention provides an isolated human nuclear receptor (designated
CC pregnane X receptor, PXR) that binds to a cytochrome P-450 mono-oxygenase
CC (CYP) promoter. The hPXR is used to identify its specific modulators,
CC and compounds that induce CYP3A4 expression (i.e. to identify drug
CC interactions, since CYP3A4 is involved in many biotransformations of
CC drugs). The modulators are potentially useful for associating particular
CC diseases and conditions with PXR and for treating such conditions.
CC Antibodies raised against hPXR can be used for determining and
CC purification of hPXR. The present sequence represents a histidine-6
CC tagged partial PXR (H1s6-PXR) sequence.

XX
S0 Sequence 316 AA;

Query Match 69.2%; Score 1585; DB 20; Length 316;
Best Local Similarity 100.0%; Pred. No. 4.3e-149;
Matches 305; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 130 SEKTGTPPLGVQGLTEEQRMIRLMDAQMKTFTTFSHKFNRLPGVLSGCELPESLQ 189
DB :|||||
DB 12 SEKTGTPPLGVQGLTEEQRMIRLMDAQMKTFTTFSHKFNRLPGVLSGCELPESLQ 71
QY 190 APSREBAKWSQVRKDCSLKVSLSLRGDSVWNTKPPADSGKEIFSLPHMADSTY 249
DB :|||||
DB 72 APSREBAKWSQVRKDCSLKVSLSLRGDSVWNTKPPADSGKEIFSLPHMADSTY 131
QY 250 MFKGIISFPAKVISYFRDLPIEDOISLKGAEFLCOLRNTYFNATGTCWEGRSYCLE 309
DB :|||||
DB 132 MFKGIISFPAKVISYFRDLPIEDOISLKGAEFLCOLRNTYFNATGTCWEGRSYCLE 191
QY 310 DTAGGFQQLLEBMLKFHYMLKKQLDHEEYVLMQALISFSPDRGVLOHRVVDQLQBOF 369
DB :|||||
DB 192 DTAGGFQQLLEBMLKFHYMLKKQLDHEEYVLMQALISFSPDRGVLOHRVVDQLQBOF 251
QY 370 ATLTKSYIECNRPQPAHRLFLKIMAMLTSLNSINAGHTORLLRIQDIHPF 429
DB :|||||
DB 252 ATLTKSYIECNRPQPAHRLFLKIMAMLTSLNSINAGHTORLLRIQDIHPF 311
QY 430 GITGS 434
DB :|||||
DB 312 gitgs 316

RESULT 13

R98521
ID R98521 standard; Protein; 386 AA.

AC R98521;

DT 14-NOV-1996 (first entry)

DE Xenopus orphan receptor 6.

XX Xenopus orphan receptor 6; XOR-6; steroid receptor; vitamin D;
XX hydroxybenzoate; mercaptohydroxybenzoate; amlinobenzoate.

OS Xenopus laevis.

PN W09622390-A1.

PD 25-JUL-1996.

PI Key Location/Qualifiers
FT Domain 37..102
FT Domain /label=DNA_binding_domain
FT Domain 183..386
FT Domain /label=Ligand_binding_domain

DR W09622390-A1.

PD 25-JUL-1996.

XX

Noted

PF 16-JAN-1996; 96WO-US00058.
 XX
 PR 17-JAN-1995; 95US-0374445.
 XX
 PA (SALK) SALK INST BIOLOGICAL STUDIES.
 XX
 PI Blumberg B, Evans RM, Umehono K;
 XX WPI; 1996-354546/35.
 DR N-PSDB; 736499.
 XX
 PT DNA encoding receptor polypeptide responsive to hydroxy, mercapto or
 PT amino benzoate(s) - useful to regulate gene transcription
 PS
 PS Claim 4; Page 27-28; 42pp; English.
 CC Xenopus orphan receptor 6 (XOR-6) (R98521) is a new member of the
 CC steroid receptor superfamily, characterised as being responsive to
 CC the presence of hydroxy, mercapto or amino benzoate(s) and as
 CC regulating the transcription of associated gene(s). It shows 73%
 CC identity in the DNA binding domain, and 42% identity in the ligand
 CC binding domain, to the human vitamin D receptor. Recombinant XOR-6
 CC can be expressed in animal cells; a cDNA clone (736499) coding for
 CC XOR-6 has been isolated. The recombinant XOR-6 may be used to
 CC regulate gene transcription or to raise antibodies of diagnostic
 CC or therapeutic appln.
 CC
 SO Sequence 366 AA;

Query Match 42.8%; Score 979.5; DB 17; Length 386;
 Best Local Similarity 48.6%; Pred. No. 6.8e-89;
 Matches 202; Conservative 60; Mismatches 105; Indels 49; Gaps 9;

QY 18 EETESVPGKPSNADDEYGVGPGICVAGDKAGYHFNWTCGCGGFFRRAMKRNARLC 77
 DB 14 EEEEDASNSCGTGEDDDGDPKICACGDRATGYHFNAMTCGCKGFFIRAVKRNLRISC 73
 QY 78 PFRKGCETTRTRROCOACRLKCLSGMKKEMTMSDEAVERALRRK-KSERGTQ 136
 DB 74 PL-qnsctvlnksnrthcgaclikcdlymrkelmsadaaveqrralrkrlkltklpt 132
 QY 137 PLGVGGLTEEQRMIRLEMDAQMTFTFSHFKNRRLPGVLSGCELPESIQAPSRREA 196
 DB 133 ppgaa-slipeqghflqtlyvghntktfdnfefsknfr-----plr--- 171
 QY 197 AKMSQVRKDLCSLKVSLQLRGDSVWNTKPPADSGGKEIFSLPHMADMTYFEKGIIS 256
 DB 172 -----rsadpt---qepqats---seafimlphladietvmikglls 207
 QY 257 FAKVISYFRLPIEDOISILKGAFFELCOLRNTVFNAETGWEGGRSLCYLEDT-AGGF 315
 DB 208 fakmlyflsldedqalixgsaevaviftnlvtfnedctnweqgftcydtedmflagf 267
 QY 316 QQLLEPMLKFRHYMLKQLLHEEYVLMQALSLFSPDPGVLOHRVVDOLOEQFAITLKS 375
 DB 268 rqlflleplvrlthmrkltnvgeeyemmaalsifasyppvcdwckiklqehalaltkd 327
 QY 376 YTECNR-POPARRFLFKITAMLTSLRSTNAOHTO--RLRIQDIDHP-ATPLMOELFG 430
 DB 328 fidsqrpsspgurlllypkimecltelrtvndhskqllelwdlqdeaplmrevfg 383

RESULT 14
 W94623
 ID W94623 standard; Protein; 423 AA.
 AC W94623;
 XX
 XX 29-APR-1999 (first entry)
 DT Rat vitamin D receptor protein VDR0.
 DE
 XX

KM Vitamin D receptor; VDR; bone density; retinoic acid derivative;
 KM steroid; bone calcium regulator; immunosuppressant; anticancer.
 XX
 XX Rattus sp.
 OS
 XX WO9856908-AL.
 PN
 PD 17-DEC-1998.
 PD
 PE 13-JUN-1997; 97WO-JP02052.
 PF
 XX 13-JUN-1997; 97WO-JP02052.
 XX
 PR 13-JUN-1997; 97WO-JP02052.
 XX
 PA (CHUS) CHUGAI SEIYAKU KK.
 PI Kato S, Ueno K;
 XX
 XX WPI; 1999-080898/07.
 DR N-PSDB; X16606.
 DR
 XX
 PT Gene for vitamin D receptor isoform protein which blocks vitamin D
 PT signal pathway - and production of recombinant isoform protein using
 PT it, for bone density assay and for screening compounds for vitamin D
 PT activity.
 PS
 PS Example 1; Fig 1; 47pp; Japanese.
 CC
 CC The present invention describes rat vitamin D receptor (VDR) isoform
 CC protein VDR1. The present sequence represents rat VDR0. VDR1 differs
 CC from the canonical form of VDR (VDR0) by lacking the amino acid sequence
 CC encoded by exon 9, but instead substitutes a short sequence
 CC (GREGREELRDLGVGDE in the rat protein) encoded by the 5'-end of intron
 CC 8. VDR1 has a dominant negative effect on the vitamin D signalling
 CC pathway. The isoform protein can be used for determining bone density,
 CC and for the screening of compounds (e.g. steroids and retinoic acid
 CC derivatives) for vitamin D activity (e.g. as bone calcium regulators,
 CC immunosuppressants or anticancer agents).
 CC
 SO Sequence 423 AA;

Query Match 35.0%; Score 801; DB 20; Length 423;
 Best Local Similarity 42.0%; Pred. No. 4e-71;
 Matches 172; Conservative 75; Mismatches 135; Indels 28; Gaps 9;

QY 38 PQCRCVGGDAATGTYHFNWTCGCGGFFRRAMKRNARLCPPRKGCETTRTRROCOAC 97
 DB 21 plicygcgdatgfhnamtccecgkgyffrrsmrkalfcpf-ngdcrltkdnrncqac 79
 QY 98 RLKRCLESGMKKEMTMSDEAVERALRRKKSERGTQPLGVGGLTEEQRMIRLEMDA 157
 DB 80 rlrctvdlgmkeflltdeevqgrrtemlrmrkeeealkslrpk-lseeghlaillda 138
 QY 158 QMKTFTFSHFNFRLPGVL--SSGCELPESIQAPSRREAQMSQVRKDLCSLKV----- 211
 DB 139 jhktdpdydfidfprrvmdsgtsysjpr-----plsfsgnsasssdllyttdlme 194
 QY 212 -----SLQLRGDSVWNTKPPADSGGKEIFSLPHMADMTYFEKGIISFAKVISYF 266
 DB 195 psqfsmldngesd-----dpsvltldleplsmplhadiavyslqkvigfakmipgfid 249
 QY 267 LPEDOISILKGAFFELCOLRNTVFNAETGWEGGRSLCYLEDT--AGGQQLLEP 323
 DB 250 ltsddqvlvixsaaalevlnmrngsftmdawdcsgdykvavtvskaghtellepl 309
 QY 324 LKFRHYMLKQLLHEEYVLMQALSLFSPDPGVLOHRVVDOLOEQFAITLKSYTECNRP 383
 DB 310 lkfgvylkrltnbeehvllmalctsvsdprrygdaklvealqdrtslnlqyictrhpp 369
 QY 384 PAHREFLFKITAMLTSLRSTNAOHTO--RLRIQDIDHP-ATPLMOELFG 430
 DB 370 pshghlyakmqlgladlrstnehsqyslsfqpensmklptlvlevfg 419

```

RESULT 15
ID W47509 standard; Protein; 423 AA.
XX AC W47509;
XX DT 26-JUN-1998 (first entry)
XX DE Rat vitamin D receptor (VDR0).
XX KW Rat; vitamin D receptor; isoform protein; VDR1; VDR0; diagnosis;
XX KW dominant negative receptor; signal transmission channel;
XX KW bone density disorder; screening.
XX OS Rattus rattus.
XX FH Key Location/Qualifiers
XX FT Misc-difference 360 /note="encoded by CAG"
XX FT W09747172-A1.
XX FN 18-DEC-1997.
XX PD 10-JUN-1997; 97MO-IB00947.
XX PE 10-JUN-1996; 96JP-0194179.
XX PR 10-JUN-1996; 96JP-0194179.
XX PA (CHUS ) CHUGAI PHARM CO LTD.
XX PA (CHUS ) CHUGAI SEIYAKU KK.
XX PI Kato S, Ueno K;
XX DR WPI; 1998-051917/05.
XX DR N-PSDB; V03129.
XX PT DNA encoding a vitamin D receptor isoform protein - useful for bone
XX PT density determination and for screening substances for vitamin D
XX PT activity
XX PS Disclosure; Fig 1; 46pp; Japanese.
XX CC A novel cDNA sequence encodes the rat vitamin D receptor isoform
XX CC protein (VDR1). The isoform differs from the normal receptor
XX CC (VDR0), which comprises the present sequence, in having the
XX CC vitamin D response element curtailed by 86 residues, and having an
XX CC extra 19 residues inserted at the C-terminal of this element. It
XX CC acts as a dominant negative receptor in the vitamin D signal
XX CC transmission channel.
XX CC The isoform protein can be used to diagnose bone density disorders,
XX CC and screen for substances having potential vitamin D-like activity.
XX SQ Sequence 423 AA;

Query Match 34.8%; Score 798; DB 19; Length 423;
Best Local Similarity 42.0%; Pred. No. 7.9e-71;
Matches 172; Conservative 74; Mismatches 136; Indels 28; Gaps 9;

QY 38 POICVCGDKATGYHFNMTCEGCKGFFRRAMKRNARLRCPEKGCETTRKTRROCOAC 97
DB 21 pricvcgdratgfhnamtcegcgkffrramkralftcpf-ngdcritkdnrrhcgac 79
QY 98 RLRKLGSGMKKEMTMSDAVEERRALIRKKKSEKRTGTPLGVOGLTEQRMTRRLMDA 157
DB 80 rlrkcvdigmhkefilldeevqirkemlmkkeealidslrpk-lseegqhliallida 138
QY 158 QMKTEDTFSHKKRNRLGVL--SSGCELPSLSIAPSREFAKMSQVAKDCSLK---- 211
DB 139 hnktydpcyadrtidrpymdgstgyspr---plstfgnsssssdlyttsldame 194
QY 212 -----SLQLRGEDGSVMYKPPADSGGKEIFSLPHEMAMSTYMKGIISPAKVIYSFRD 266

```

```

DB 195 psgfsmldngesd-----dpesvldleplsmldpladlvysiqkylgfkamipgrfd 249
QY 267 LPRIEDISLLKGAAPFLCOLRENTVFNMAETGWEGRSLSYCLEDR---AGGRQULLLEPM 323
DB 250 ltsddqivllksalevlnlrngsfndmawdcsgdqxydvldvskaghtllellep1 309
QY 324 LKPFHYMLKKLOJHEEYVLMQALISLSPDRPGVLQHRVVDQLQEQFAITLKSIECNRPQ 383
DB 310 ikfvgvllklnlheeenvlmaiclvspdrpvgdakivealqdrlnltlgtfyrcthrpp 369
QY 384 PAHRFLFKIMAMLTSLRSINAOHTQ--RLLRIDIHFP-ATPLMQELFEG 430
DB 370 pgsnqlyakmiglkladrlslnesksqyslsfqpensmkltpvlvlefvf 419

```

Search completed: February 18, 2001, 14:29:39
Job time: 8216 sec

Db 11 CVCGDGATGYHNNALTCBCKGFFRRVYSKISGPTCEP-AGSCVSKTQRHRCPCARLQ 69
 QY 101 KCLESQMKKIMSDAEVBERRALIRKKSERTGTPLGVQGLTEORMMIRELMDAOMK 160
 Db 70 KCLDAGMRKMDILSAELALBRKAQARRAQOTPVQ-----LSKQEBLITLLGAHTR 123
 QY 161 TPDTSSEHFNRLPGVLSGCELPESLQAPSEBAKMSQVRKDLCSLKSVLQRGEDG 220
 Db 124 HMGTFEQFQVFRPAHLFIHQ-PLPTLAP----- 153
 QY 221 SWNNKRPADSGKEIFSLPHMADSTYMEKGIISFAKVISYFRDLPTEQISILKGA 280
 Db 154 -----VPLTHFADINTFMVLQVIFKFDLPVFRSLPLEQISILKGA 198
 QY 281 FELCOLRNTVFNAGTWEBCGRLSYCLBDTAG-GFOQLLEPMLKFHYMLKRLQHEE 339
 Db 199 VEICHVYANTTFCLOTQNFCLGPRKRTIEDGARVGFVEFLLELHFHGTLLKRLQLOPE 258
 QY 340 YVLMQALISFSPDRPGVLOHRRVVDLQEQFAITLKSIECNRPQPAHRELFKIMAMTE 399
 Db 259 YVLLAAMALFSPDRPGVLOHRRVVDLQEQFAITLKSIECNRPQPAHRELFKIMAMTE 318
 QY 400 LRSINAGHQRRLRIODIHFPATPLMOEL 428
 Db 319 LRSINEAYGYQIOHIGLSAM-MPLLOEI 346

RESULT 2

US-08-458-686-10
 ; Sequence 10, Application US/08458686
 ; Patent No. 5710017

GENERAL INFORMATION:

APPLICANT: David D. Moore et al.
 TITLE OF INVENTION: CAR RECEPTORS AND RELATED
 NUMBER OF SEQUENCES: 10
 CORRESPONDENCE ADDRESS:

ADDRESSEE: Fish & Richardson
 STREET: 225 Franklin Street
 CITY: Boston

STATE: Massachusetts
 COUNTRY: U.S.A.

ZIP: 02110-2804

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5" Diskette, 1.44 Mb

COMPUTER: IBM PS/2 Model 502 or 55SX

OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)

SOFTWARE: WordPerfect (Version 5.0)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/458,686

FILING DATE:

CLASSIFICATION: 514

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/843,350

FILING DATE: February 26, 1992

ATTORNEY/AGENT INFORMATION:

NAME: Paul T. Clark

REGISTRATION NUMBER: 30,162

REFERENCE/DOCKET NUMBER: 00786/126001

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 542-5070

TELEFAX: (617) 542-8906

TELEX: 200154

INFORMATION FOR SEQ ID NO: 10:

SEQUENCE CHARACTERISTICS:

LENGTH: 348

TYPE: amino acid

STRANDEDNESS: N/A

TOPOLOGY: linear

US-08-458-686-10

Best Local Similarity 42.2%; Pred. No. 4,1e-68;
 Matches 164; Conservative 62; Mismatches 109; Indels 54; Gaps 6;
 QY 41 CVCGDGATGYHNNALTCBCKGFFRRVYSKISGPTCEP-AGSCVSKTQRHRCPCARLQ 100
 Db 11 CVCGDGATGYHNNALTCBCKGFFRRVYSKISGPTCEP-AGSCVSKTQRHRCPCARLQ 69
 QY 101 KCLESQMKKIMSDAEVBERRALIRKKSERTGTPLGVQGLTEORMMIRELMDAOMK 160
 Db 70 KCLDAGMRKMDILSAELALBRKAQARRAQOTPVQ-----LSKQEBLITLLGAHTR 123
 QY 161 TPDTSSEHFNRLPGVLSGCELPESLQAPSEBAKMSQVRKDLCSLKSVLQRGEDG 220
 Db 124 HMGTFEQFQVFRPAHLFIHQ-PLPTLAP----- 153
 QY 221 SWNNKRPADSGKEIFSLPHMADSTYMEKGIISFAKVISYFRDLPTEQISILKGA 280
 Db 154 -----VPLTHFADINTFMVLQVIFKFDLPVFRSLPLEQISILKGA 198
 QY 281 FELCOLRNTVFNAGTWEBCGRLSYCLBDTAG-GFOQLLEPMLKFHYMLKRLQHEE 339
 Db 199 VEICHVYANTTFCLOTQNFCLGPRKRTIEDGARVGFVEFLLELHFHGTLLKRLQLOPE 258
 QY 340 YVLMQALISFSPDRPGVLOHRRVVDLQEQFAITLKSIECNRPQPAHRELFKIMAMTE 399
 Db 259 YVLLAAMALFSPDRPGVLOHRRVVDLQEQFAITLKSIECNRPQPAHRELFKIMAMTE 318
 QY 400 LRSINAGHQRRLRIODIHFPATPLMOEL 428
 Db 319 LRSINEAYGYQIOHIGLSAM-MPLLOEI 346

RESULT 3

US-07-843-350C-10
 ; Sequence 10, Application US/07843350C
 ; Patent No. 5756448

GENERAL INFORMATION:

APPLICANT: David D. Moore et al.
 TITLE OF INVENTION: CAR RECEPTORS AND RELATED
 NUMBER OF SEQUENCES: 10
 CORRESPONDENCE ADDRESS:

ADDRESSEE: Fish & Richardson
 STREET: 225 Franklin Street
 CITY: Boston

STATE: Massachusetts
 COUNTRY: U.S.A.

ZIP: 02110-2804

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5" Diskette, 1.44 Mb

COMPUTER: IBM PS/2 Model 502 or 55SX

OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)

SOFTWARE: WordPerfect (Version 5.0)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/843,350C

FILING DATE: February 26, 1992

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Paul T. Clark

REGISTRATION NUMBER: 30,162

REFERENCE/DOCKET NUMBER: 00786/126001

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 542-5070

TELEFAX: (617) 542-8906

TELEX: 200154

INFORMATION FOR SEQ ID NO: 10:

SEQUENCE CHARACTERISTICS:

LENGTH: 348

TYPE: amino acid

STRANDEDNESS: N/A

US-07-843-350C-10

Query Match

33.5%; Score 768; DB 1; Length 348;


```

? ZIP: 07065-0907
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patentin Release #1.0, Version #1.25
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/330,518
? FILING DATE:
? CLASSIFICATION: 435
? ATTORNEY/AGENT INFORMATION:
? NAME: Dolan, Catherine A.
? REGISTRATION NUMBER: 36,502
? REFERENCE/DOCKET NUMBER: 19316
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (908) 594-4283
? TELEFAX: (908) 594-4720
? INFORMATION FOR SEQ ID NO: 2:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 461 amino acids
? TYPE: amino acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: peptide
? HYPOTHEICAL: NO
? ANTI-SENSE: NO
? US-08-330-518-2

```

```

Query Match      21.7%; Score 498; DB 1; Length 461;
Best Local Similarity 29.0%; Pred. No. 3.6e-41;
Matches 126; Conservative 86; Mismatches 125; Indels 98; Gaps 14;

```

```

QY 37 GPOICRYCGDKATGYHFNWTCGCGFFRRAMKRNARLCFPR-KGACEITRRKRCQ 95
DB 83 GHELCRYCGDKASGFHYNVSCGCGFFRRSVRGARRVACRGGTCOMDAFMRKCQ 142
QY 96 ACRLKRCLESGMKKEMIMSDAEVEERRALIKRKSERTGTQ--PLGVQ----- 142
DB 143 QCRLRCKEAGMRQCVLSEQIRKK--IRKQOQESQSOSQSPVGGSSSSASGPGA 200
QY 143 -----LTEQRMMIRELMDAOKTFTDTSFHKFNRLPGVLSGCE 183
DB 201 SPGSENGSQSGEGEGVQLTAQELMIQOLVAQLOCNKRSF----- 244
QY 184 LPESIQAPSEBAKMSQVRKDCSLKYSIQLRGEGSVWNKKPPADSGKELFSILPHM 243
DB 245 -----DQKVTWMP-----LGAD-----PQSRDAQQRFRA--HF 271
QY 244 ADMSTYFKGIISFAVVISYFRDLPIEDQISLLKGAFFELCOLRNTVFNATGTWEC-- 301
DB 272 TELAIISVQELVDFAKQVGFQLGRBDIALKASTIEIMLEFARRYNHET---ECIT 328
QY 302 --GRLSYCLEDD-TAGGFQOLLEPMLEKHYMKKLIQHEEYVLMQAISLSPDRPGVLQ 358
DB 329 FLKDFYSDDDHFRAGLQVEFINPIFEFSRARRRLGLDAEVALIAINISADPNVQE 388
QY 359 HRVVDQLOEPAITLTKSYIECNRPQAHFELFKIMAMTELRSINAQHTQL--LRIOD 416
DB 389 PGRVVALQOPEYVALLSTYRIKRPQDQLR--FPRMLKLVSLRTLSVSHSEQVALLRLQD 446
QY 417 IHPFATPLMQELFGI 431
DB 447 --KLLPPLLSIEIMDV 459

```

```

RESULT 6
US-08-330-283-2
; Sequence No., Application US/08330283
; Patent No. 5679518
; GENERAL INFORMATION:
; APPLICANT: Friedman, Eitan
; APPLICANT: Holloway, M. Katharine

```

```

? APPLICANT: Rodan, Gideon
? APPLICANT: Rutledge, Su Jane
? APPLICANT: Schmidt, Azriel
? APPLICANT: Vogel, Robert
? TITLE OF INVENTION: METHOD FOR FINDING RECEPTOR POTENTIATORS
? NUMBER OF SEQUENCES: 5
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Merck & Co., Inc.
? STREET: 126 East Lincoln Avenue
? CITY: Rahway
? STATE: New Jersey
? COUNTRY: US
? ZIP: 07065-0907
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patentin Release #1.0, Version #1.25
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/330,283
? FILING DATE:
? CLASSIFICATION:
? ATTORNEY/AGENT INFORMATION:
? NAME: Dolan, Catherine A.
? REGISTRATION NUMBER: 36,502
? REFERENCE/DOCKET NUMBER: 19327
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (908) 594-4283
? TELEFAX: (908) 594-4720
? INFORMATION FOR SEQ ID NO: 2:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 461 amino acids
? TYPE: amino acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: peptide
? HYPOTHEICAL: NO
? ANTI-SENSE: NO
? US-08-330-283-2

```

```

Query Match      21.7%; Score 498; DB 1; Length 461;
Best Local Similarity 29.0%; Pred. No. 3.6e-41;
Matches 126; Conservative 86; Mismatches 125; Indels 98; Gaps 14;

```

```

QY 37 GPOICRYCGDKATGYHFNWTCGCGFFRRAMKRNARLCFPR-KGACEITRRKRCQ 95
DB 83 GHELCRYCGDKASGFHYNVSCGCGFFRRSVRGARRVACRGGTCOMDAFMRKCQ 142
QY 96 ACRLKRCLESGMKKEMIMSDAEVEERRALIKRKSERTGTQ--PLGVQ----- 142
DB 143 QCRLRCKEAGMRQCVLSEQIRKK--IRKQOQESQSOSQSPVGGSSSSASGPGA 200
QY 143 -----LTEQRMMIRELMDAOKTFTDTSFHKFNRLPGVLSGCE 183
DB 201 SPGSENGSQSGEGEGVQLTAQELMIQOLVAQLOCNKRSF----- 244
QY 184 LPESIQAPSEBAKMSQVRKDCSLKYSIQLRGEGSVWNKKPPADSGKELFSILPHM 243
DB 245 -----DQKVTWMP-----LGAD-----PQSRDAQQRFRA--HF 271
QY 244 ADMSTYFKGIISFAVVISYFRDLPIEDQISLLKGAFFELCOLRNTVFNATGTWEC-- 301
DB 272 TELAIISVQELVDFAKQVGFQLGRBDIALKASTIEIMLEFARRYNHET---ECIT 328
QY 302 --GRLSYCLEDD-TAGGFQOLLEPMLEKHYMKKLIQHEEYVLMQAISLSPDRPGVLQ 358
DB 329 FLKDFYSDDDHFRAGLQVEFINPIFEFSRARRRLGLDAEVALIAINISADPNVQE 388
QY 359 HRVVDQLOEPAITLTKSYIECNRPQAHFELFKIMAMTELRSINAQHTQL--LRIOD 416
DB 389 PGRVVALQOPEYVALLSTYRIKRPQDQLR--FPRMLKLVSLRTLSVSHSEQVALLRLQD 446

```


Db 83 GHELCRCVCGDKRAGFHYNYLSCGCGKGFRRSVRGARVACRGCGTCCQDAFMRRCQ 142
 QY 96 ACRLKRCLESGMKKEMINSDEAVERRALIKRKSERTGTQ---PLGVQG----- 142
 Db 143 QCRLRCKEAGMRKQCVLSEEDIRKK--IRKQOQESQSQSVGVPGSSSSASGPGA 200
 QY 143 -----LTEQRMMIRELMDAQMTFTDTSFHKFRPLPGVLSGCE 183
 Db 201 SPGSEAGSQSGSEBGVQLTAQELMIQQLVAQLOQCKRSFS----- 244
 QY 184 LPESLAPSRERAAKWSQVRKDLCSLKVSLQLRGDSVWYKPPADSGGKTEFSLPHM 243
 Db 245 -----DQPVTPWP-----LGAD-----PQSRARQORFA---HF 271
 QY 244 ADMSTYMGKIISEFAKVISYFRDLPEDQISLKGAFELQLRNTVFNATGTWEC-- 301
 Db 272 TELATISVOEYDFAKQVGFLOLGRDQIALKASTEIMLETARRYNHET---ECIT 328
 QY 302 --GRLSYCLD--TAGGFQOLLLEPMLEKHYMLKQLHEEYVLMQALSLFSPDRPGVLO 358
 Db 329 FLKDFYSKDDFHRAGLQVEFINPFEFSRAMRRLGLDADAYALLAINITSADRPVQE 388
 QY 359 HRVVDQLOEQFALTKSTIECNRPQARHFLKIMAMLTLSRINAOHTQRL--LRID 416
 Db 389 PGRVEALQOPYEVALSTYRIKRPQDLR--FPRMLKVLSTLTSVSHSQVAFALRLD 446
 QY 417 IHPPATPLMOELFGI 431
 Db 447 --KKLPPLSEIMDV 459

RESULT 9

PCT-US95-13931-2
 Sequence 2, Application PC/TUS9513931

GENERAL INFORMATION:

APPLICANT: Friedmann, Eitan

APPLICANT: Holloway, M. Katharine

APPLICANT: Rodan, Gideon

APPLICANT: Schmidt, Azriel

APPLICANT: Vogel, Robert

TITLE OF INVENTION: USE OF RECEPTOR POTENTIATORS

NUMBER OF SEQUENCES: 5

CORRESPONDENCE ADDRESS:

ADDRESS: Merck & Co., Inc.

STREET: 126 East Lincoln Avenue

CITY: Rahway

STATE: New Jersey

COUNTRY: US

ZIP: 07065-0907

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US95/13931

FILING DATE:

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Quaglinato, Carol S.

REGISTRATION NUMBER: 35,330

REFERENCE/DOCKET NUMBER: 19316 PCT

TELECOMMUNICATION INFORMATION:

TELEPHONE: (908) 594-3809

TELEFAX: (908) 594-4720

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 461 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

HYPOTHETICAL: NO

ANTI-SENSE: NO
 PCT-US95-13931-2

Query Match 21.7%; Score 498; DB 4; Length 461;
 Best Local Similarity 29.0%; Pred. No. 3,6e-41;
 Matches 126; Conservative 86; Mismatches 125; Indels 98; Gaps 14;

QY 37 GPQICRYCGDKATGTFHYNNWTCBGCCKGFRRAMKRNALRCFPR-KGACETRTKTRCQ 95
 Db 83 GHELCRCVCGDKRAGFHYNYLSCGCGKGFRRSVRGARVACRGCGTCCQDAFMRRCQ 142
 QY 96 ACRLKRCLESGMKKEMINSDEAVERRALIKRKSERTGTQ---PLGVQG----- 142
 Db 143 QCRLRCKEAGMRKQCVLSEEDIRKK--IRKQOQESQSQSVGVPGSSSSASGPGA 200
 QY 143 -----LTEQRMMIRELMDAQMTFTDTSFHKFRPLPGVLSGCE 183
 Db 201 SPGSEAGSQSGSEBGVQLTAQELMIQQLVAQLOQCKRSFS----- 244
 QY 184 LPESLAPSRERAAKWSQVRKDLCSLKVSLQLRGDSVWYKPPADSGGKTEFSLPHM 243
 Db 245 -----DQPVTPWP-----LGAD-----PQSRARQORFA---HF 271
 QY 244 ADMSTYMGKIISEFAKVISYFRDLPEDQISLKGAFELQLRNTVFNATGTWEC-- 301
 Db 272 TELATISVOEYDFAKQVGFLOLGRDQIALKASTEIMLETARRYNHET---ECIT 328
 QY 302 --GRLSYCLD--TAGGFQOLLLEPMLEKHYMLKQLHEEYVLMQALSLFSPDRPGVLO 358
 Db 329 FLKDFYSKDDFHRAGLQVEFINPFEFSRAMRRLGLDADAYALLAINITSADRPVQE 388
 QY 359 HRVVDQLOEQFALTKSTIECNRPQARHFLKIMAMLTLSRINAOHTQRL--LRID 416
 Db 389 PGRVEALQOPYEVALSTYRIKRPQDLR--FPRMLKVLSTLTSVSHSQVAFALRLD 446
 QY 417 IHPPATPLMOELFGI 431
 Db 447 --KKLPPLSEIMDV 459

RESULT 10

US-08-342-411A-2
 Sequence 2, Application US/08342411A

Patent No. 5639616

GENERAL INFORMATION:

APPLICANT: LIAO, Shutsung

APPLICANT: SONG, Ching

TITLE OF INVENTION: UNUSUAL NUCLEAR RECEPTOR:

TITLE OF INVENTION: COMPOSITIONS AND METHODS

NUMBER OF SEQUENCES: 38

CORRESPONDENCE ADDRESS:

ADDRESS: Arnold, White & Durkee

STREET: P.O. Box 4433

CITY: Houston

STATE: TX

COUNTRY: USA

ZIP: 77210-4433

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/342,411A

FILING DATE: 18-NOV-1994

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: KITCHELL, BARBARA S.

REGISTRATION NUMBER: 33,928

REFERENCE/DOCKET NUMBER: ARCD154

TELECOMMUNICATION INFORMATION:

TELEPHONE: (512) 418-3000

TELEFAX: (713) 789-2679
 TELETYPE: 79-0924
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 460 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 US-08-342-411A-2

Query Match 21.7%: Score 496; DB 1; Length 460;
 Best Local Similarity 29.0%; Pred. No. 5,6e-41;
 Matches 126; Conservative 86; Mismatches 125; Indels 98; Gaps 14;

37 GPQIRVGGDKATGHNVTGCEGKGFRRMRKRNALRCPFR-KGACETIRKTRRQCQ 95
 82 GHELCRVGGDKASGHVNVLDGCKGFFRSVVGARRACRGGGCGQADAFRRKQCQ 141
 96 ACRLKRCLESGMKKMSDEAVEERALLIRKKSERTGTQ---PLGVQ----- 142
 142 QCRLRKCKEAGRECVLSEEDIRKK--IRKQGGGQSGQSGSPVGPSSSSASGPGA 199
 143 -----LIEQRMIRLEMDAQMTFTTFSHFNKFRPLRPLVLSGCE 183
 200 SPGSGEASGSGSGEGVQLTMAQELMTQVLAQLOCNKRSFS----- 243
 184 LPESLQASREBAKMSQVRDCLSKVSLQRGEDGSVMYKRPADSGKEISLPHM 243
 244 -----DQPVTPWP-----LGAD-----PSRDARQQRPA--HF 270
 244 ADMSTYMGKISFAKVISYFNDLPIDQISLKGAFELCOLRNTVFNAETGWECC 301
 271 TELAIISQVEIDFKVQPGFQLGREDQIALKASTIEIMLETRARVNHET--ECIT 327
 302 --GRUSTYLED-TAGCFQQLLEPKLFRHYMKKQLQHEEYVLMQALSLSPDRPGVLO 358
 328 FLKDFPTYSKDDHFRAGLQVEFINPIFEFSRMRRLGLDAEYALVLAITSADRPVQE 387
 359 HNVQDLQEQFATLKYSTECHNRQPAHRELFKIMAMLETLSINAOHQRL--LRID 416
 388 PERVELDQPYEALSTYRKRPODLR--FPRMLKLVSLRTLVSVHSEQVAFRLRD 445
 417 IHPFATPLMQELFGI 431
 446 --KKLPPLSEIWDV 458

RESULT 11
 US-08-333-358-8
 Sequence 8, Application US/08333358
 Patent No. 5571696
 GENERAL INFORMATION:
 APPLICANT: EVANS Ph.D., RONALD M.
 APPLICANT: MANGELSDORF Ph.D., DAVID J.
 APPLICANT: ONG Ms., ESTELITA S.
 APPLICANT: ONG Ms., ANTHONY E.
 APPLICANT: BORGMEYER Ph.D., VINCENT K.
 APPLICANT: GIGUERE Ph.D., VINCENT NMN
 APPLICANT: YAO Mr., TSO-PANG NMN
 TITLE OF INVENTION: NOVEL RECEPTORS
 NUMBER OF SEQUENCES: 14
 CORRESPONDENCE ADDRESS:
 ADDRESS: Pretty, Schroeder, Brueggemann & Clark
 STREET: 444 So. Flower St., Suite 2000
 CITY: Los Angeles
 STATE: CA
 COUNTRY: US
 ZIP: 90071-2921
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/333,358
 FILING DATE:
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/07/761,068
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: Reiter Ph.D., Stephen E.
 REGISTRATION NUMBER: 31192
 REFERENCE/DOCKET NUMBER: P31 8936
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (619) 535-9001
 TELEFAX: (619) 535-8949
 INFORMATION FOR SEQ ID NO: 8:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 440 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-333-358-8

Query Match 21.5%: Score 491.5; DB 1; Length 440;
 Best Local Similarity 29.7%; Pred. No. 1.5e-40;
 Matches 127; Conservative 85; Mismatches 145; Indels 71; Gaps 12;

18 EDTESVYGRKSVNADEVEGPGQICRVGDKATGHNVTGCEGKGFRRMRKRNALRLC 77
 68 EPTEIRPQKRRKPPAPKMLGNELCYGDKASGFHNVLSCEGCKGFFRSYKGAHYIC 127
 78 PFRGACEIRKTRROCOACRLKRCLESGMKKMSDEAVEERALLIRKKS----- 131
 128 -HSGGHPMDTYRRKQCEBRLKRCQAGRECVLSEEDIRKK--LKRQEEQAHNS 184
 132 ---RTGTPLVGQGLTEQRMIRLEMDAQMTFTTFSHFNKFRPLRPLVLSGCELPESL 188
 185 LPPRRSSPPQILPOLSPQGMIEKLVAAQOCNRRSFS----- 223
 189 QASREBAKMSQVRDCLSKVSLQRGEDGSVMYKRPADSGKEISLPHMADMT 248
 224 ---DRLVTPWPNAP-----DPRSBRARQQRPA--HFTELAI 255
 249 YMEKGISFAKVISYFNDLPIDQISLKGAFELCOLRNTVFNAETGWECCRLSY 306
 256 VSVQVEIDFKVQPGFQLGREDQIALKASTIEIMLETRARVNHET--LNDPSY 314
 307 CLEDTA-GGFQQLLEPKLFRHYMKKQLQHEEYVLMQALSLSPDRPGVLOHRYVDOL 365
 315 NREDFAKAGLQVEFINPIFEFSRMRRLGLDAEYALVLAITSADRPVQDLQVERL 374
 366 QEQFATLKYSTECHNRQPAHRELFKIMAMLETLSINAOHQRL--LRIDHFFATP 423
 375 QHVEYVLAHVAISIHDP--DRLMFPRMLKLVSLRTLVSVHSEQVAFRLRD--KKLP 430
 424 IMQELFGI 431
 431 LLSIWDV 438

RESULT 12
 US-08-463-694-8
 Sequence 8, Application US/08463694
 Patent No. 5696233
 GENERAL INFORMATION:
 APPLICANT: EVANS Ph.D., RONALD M.
 APPLICANT: MANGELSDORF Ph.D., DAVID J.
 APPLICANT: ONG Ms., ESTELITA S.
 APPLICANT: ONG Ms., ANTHONY E.
 APPLICANT: BORGMEYER Ph.D., VINCENT K.
 APPLICANT: GIGUERE Ph.D., VINCENT NMN
 APPLICANT: YAO Mr., TSO-PANG NMN
 TITLE OF INVENTION: NOVEL RECEPTORS

```

NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESS: Pretty, Schroeder, Brueggemann & Clark
STREET: 444 So. Flower St., Suite 2000
CITY: Los Angeles
STATE: CA
COUNTRY: US
ZIP: 90071-2921
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/463,694
FILING DATE: 05-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/761,068
FILING DATE: 17-SEP-1991
STATE: CA
COUNTRY: US
ZIP: 90071-2921
ATTORNEY/AGENT INFORMATION:
NAME: Reiter Ph.D., Stephen E.
REGISTRATION NUMBER: 31192
REFERENCE/DOCKET NUMBER: P31 8936
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 440 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-463-694-8

```

Query Match 21.5%; Score 491.5; DB 1; Length 440;
Best Local Similarity 29.7%; Pred. No. 1.5e-40;
Matches 127; Conservative 85; Mismatches 145; Indels 71; Gaps 12;

18 EDTSVGPKPSVNADEVEGPOICRYCGDKATGYHFNWVTCGCGKGFRRRAKRNARLRC 77
| | | | | : : : | | | | | : : : | | | | | : : : | | | | | : : : |
68 EPTETRPQKRRKGPAPKMGNEICSYCGDKASGFHYNVLSCGCGKGFRRRSYIKGAHYIC 127
| | | | | : : : | | | | | : : : | | | | | : : : | | | | | : : : |
78 PFRKACEITRTTRROCAQLRKCLSGMKKEMIMSDAVERALIKRKSE----- 131
| | | | | : : : | | | | | : : : | | | | | : : : | | | | | : : : |
128 -HSGGCHPMDYMRKRCQRLKRCQAGMRBEVLSGEQIRLKK--LKRQEEQAHATS 184
| | | | | : : : | | | | | : : : | | | | | : : : | | | | | : : : |
132 ---RTGTPLYVQGLTEQRMMIRELMDQMTFTTFSHKFNFLPGVLSGCELPESL 188
| | | | | : : : | | | | | : : : | | | | | : : : | | | | | : : : |
185 LPPRRSSPOLLPOLSPQGLGMIETLVAAQOCNRRSF----- 223
| | | | | : : : | | | | | : : : | | | | | : : : | | | | | : : : |
189 QAPSRSEAKSKQVRKDCSLKVSQLRGEGDSVNNYKPPADSGGKEIFSLPHMADMT 248
| | | | | : : : | | | | | : : : | | | | | : : : | | | | | : : : |
224 ---DLRTVPMPAP-----DPHSREARQORFA---HFTETLAI 255
| | | | | : : : | | | | | : : : | | | | | : : : | | | | | : : : |
249 YMFKGIISFAKVISYFRDLPIDQISLKGAFELCOLRFTVFN--AETGWECCGRLSY 306
| | | | | : : : | | | | | : : : | | | | | : : : | | | | | : : : |
256 VSVQITVFAQLPQGLDLSREDQIALKTSALETSMLETSTRYRNPGESEITF--LKDFSY 314
| | | | | : : : | | | | | : : : | | | | | : : : | | | | | : : : |
307 CLEDTA--GGFOOLLLEPMKFMMLKQLHEEYVLMQAISLSPSPDPGVLAQHRVVDOL 365
| | | | | : : : | | | | | : : : | | | | | : : : | | | | | : : : |
315 NREDPAKAGLOVEFINPFIFFSRAMNELQLNDAEPALLIAISISADRPVNOQLOVERL 374
| | | | | : : : | | | | | : : : | | | | | : : : | | | | | : : : |
366 QEQFAITLKSYIECNRPQAPHRFLFKITAMULTELRSINAOHTORL--LRIQDIHPATP 423
| | | | | : : : | | | | | : : : | | | | | : : : | | | | | : : : |
375 QHTYEALHAAYSIHHPH--DRIMEPRMIMKLVSLRTLSVSHQVYALRLQD--KKLP 430
| | | | | : : : | | | | | : : : | | | | | : : : | | | | | : : : |
424 LMOELFGI 431
| | | | | : : : | | | | | : : : | | | | | : : : | | | | | : : : |
431 LLSLITWDV 438
| | | | | : : : | | | | | : : : | | | | | : : : | | | | | : : : |

```

RESULT 13
US-08-694-501-8
Sequence 8, Application US/08694501
Patent No. 5710004
GENERAL INFORMATION:
APPLICANT: EVANS Ph.D., RONALD M.
APPLICANT: MANGELSDORF Ph.D., DAVID J.
APPLICANT: ONG MS., ESTELITA S.
APPLICANT: ONG Ph.D., ANTHONY E.
APPLICANT: BORGMEYER Ph.D., UWE K.
APPLICANT: GIBBERE Ph.D., VINCENT NAM
APPLICANT: YAO Mr., TSO-PANG NAM
TITLE OF INVENTION: NOVEL RECEPTORS
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESS: Pretty, Schroeder, Brueggemann & Clark
STREET: 444 So. Flower St., Suite 2000
CITY: Los Angeles
STATE: CA
COUNTRY: US
ZIP: 90071-2921
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/694,501
FILING DATE: 07-AUG-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/333,358
FILING DATE:
APPLICATION NUMBER: US/07/761,068
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Reiter Ph.D., Stephen E.
REGISTRATION NUMBER: 31192
REFERENCE/DOCKET NUMBER: P31 8936
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 440 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-694-501-8

```

Query Match 21.5%; Score 491.5; DB 1; Length 440;
Best Local Similarity 29.7%; Pred. No. 1.5e-40;
Matches 127; Conservative 85; Mismatches 145; Indels 71; Gaps 12;

18 EDTSVGPKPSVNADEVEGPOICRYCGDKATGYHFNWVTCGCGKGFRRRAKRNARLRC 77
| | | | | : : : | | | | | : : : | | | | | : : : | | | | | : : : |
68 EPTETRPQKRRKGPAPKMGNEICSYCGDKASGFHYNVLSCGCGKGFRRRSYIKGAHYIC 127
| | | | | : : : | | | | | : : : | | | | | : : : | | | | | : : : |
78 PFRKACEITRTTRROCAQLRKCLSGMKKEMIMSDAVERALIKRKSE----- 131
| | | | | : : : | | | | | : : : | | | | | : : : | | | | | : : : |
128 -HSGGCHPMDYMRKRCQRLKRCQAGMRBEVLSGEQIRLKK--LKRQEEQAHATS 184
| | | | | : : : | | | | | : : : | | | | | : : : | | | | | : : : |
132 ---RTGTPLYVQGLTEQRMMIRELMDQMTFTTFSHKFNFLPGVLSGCELPESL 188
| | | | | : : : | | | | | : : : | | | | | : : : | | | | | : : : |
185 LPPRRSSPOLLPOLSPQGLGMIETLVAAQOCNRRSF----- 223
| | | | | : : : | | | | | : : : | | | | | : : : | | | | | : : : |
189 QAPSRSEAKSKQVRKDCSLKVSQLRGEGDSVNNYKPPADSGGKEIFSLPHMADMT 248
| | | | | : : : | | | | | : : : | | | | | : : : | | | | | : : : |
224 ---DLRTVPMPAP-----DPHSREARQORFA---HFTETLAI 255
| | | | | : : : | | | | | : : : | | | | | : : : | | | | | : : : |
249 YMFKGIISFAKVISYFRDLPIDQISLKGAFELCOLRFTVFN--AETGWECCGRLSY 306
| | | | | : : : | | | | | : : : | | | | | : : : | | | | | : : : |

Db 256 VSOEIVDPKQPLGFLQSLREDOIALKLTSAIEVMLLETSRRYNGSEITF-LKDFSY 314
QY 307 CLEDTA-GGFQOLLEPMLEKFLKQLHEEYVLMQATISLSPDRGCVLQHNVPOL 365
Db 315 NREDFAKAGLOVEFINIFEFSSRAMELOLNDAEFALLAISISADRNVDOLQOVERL 374
QY 366 QEOFAITLKSIECNRPQAPARFLFKTMAMTELRSINAOHTOL--LRIDIHPPATP 423
Db 375 QHTVEALAHAYVSIHHPH--DRIMEPRLMKLVSLRTLSVSHQVFLRLQD--KKLP 430
QY 424 LMOELFGI 431
Db 431 LLEIMDV 438

RESULT 14
US-08-373-935-1
; Sequence 1, Application US/08373935
; Patent No. 5747661
; GENERAL INFORMATION:
; APPLICANT: Evans, Ronald M.
; APPLICANT: Mangelsdorf, David J.
; APPLICANT: Willy, Patricia J.
; TITLE OF INVENTION: IDENTIFICATION OF A DISTINCT
; TITLE OF INVENTION: RETINOID-RESPONSIVE PATHWAY AND USES THEREFOR
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pretty, Schroeder, Brueggemann & Clark
; STREET: 444 South Flower Street, Suite 2000
; CITY: Los Angeles
; STATE: CA
; COUNTRY: USA
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/373.935
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Reiter, Stephen E.
; REGISTRATION NUMBER: 31,192
; REFERENCE/DOCKET NUMBER: P41 9894
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-546-4737
; TELEFAX: 619-546-4737
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 447 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-373-935-1

Query Match 21.5%; Score 491.5; DB 1; Length 447;
Best Local Similarity 29.7%; Pred. No. 1.5e-40;
Matches 127; Conservative 85; Mismatches 145; Indels 71; Gaps 12;

QY 18 EDTESVPRKPSVNDDEYGGQICVCDKATGYHFNVTCEGCKGFFRRAMRNALRC 77
Db 75 EPTETRPQRRKKGPAKPKLGNELSYVCGDKASGPHYVNLSCGCKGFFRRSVIKGAYIC 134
QY 78 PFRKACETTRKTRROCCARLRCLESGMKKEMIMSDAEVEERBALIKRKE----- 131
Db 135 -HSGHCHMDIYMRKCECLTRCRQAGAREEVLSEBOIRLKK--LKRQEEQAHAYS 191
QY 132 ---RGTGPLVQGLTEQQRMMRELMDAOKMTDTTFSHKFNRLPGVLSGCELPESL 188
Db 192 LPRRSPPQILPOLSPOLGMIETLVAAQOCNRSSFS----- 230

QY 189 QAPSRERANWSQVRKDLCSLVSLQNGEDGSVNNYRPPADSGKEITSLPHMADMT 248
Db 231 ---DRRYTFMPWAP-----DPHSREARQORFA---HFTELAI 262
QY 249 YMRGIISFAKIVSYPRDLPIEDQISILKGAFFELCOLRNTVFN--AETGTWECRLSY 306
Db 263 VSOEIVDPKQPLGFLQSLREDOIALKLTSAIEVMLLETSRRYNGSEITF-LKDFSY 321
QY 307 CLEDTA-GGFQOLLEPMLEKFLKQLHEEYVLMQATISLSPDRGCVLQHNVPOL 365
Db 322 NREDFAKAGLOVEFINIFEFSSRAMELOLNDAEFALLAISISADRNVDOLQOVERL 381
QY 366 QEOFAITLKSIECNRPQAPARFLFKTMAMTELRSINAOHTOL--LRIDIHPPATP 423
Db 382 QHTVEALAHAYVSIHHPH--DRIMEPRLMKLVSLRTLSVSHQVFLRLQD--KKLP 437
QY 424 LMOELFGI 431
Db 438 LLEIMDV 445

RESULT 15
US-08-372-652-3
; Sequence 3, Application US/08372652
; Patent No. 5932699
; GENERAL INFORMATION:
; APPLICANT: Moore, David
; APPLICANT: Secl, Wong
; APPLICANT: Choi, Hwang-Sik
; TITLE OF INVENTION: RETINOID X RECEPTOR-INTERACTING
; TITLE OF INVENTION: POLYPEPTIDES AND RELATED MOLECULES AND METHODS
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street, Suite 3100
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/372.652
; FILING DATE: 13-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Clark, Paul T.
; REGISTRATION NUMBER: 30,162
; REFERENCE/DOCKET NUMBER: 00786/246001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 446 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-372-652-3

Query Match 21.4%; Score 490; DB 2; Length 446;
Best Local Similarity 29.1%; Pred. No. 2.1e-40;
Matches 125; Conservative 84; Mismatches 128; Indels 92; Gaps 14;

QY 37 GPQICRYGKATGYHFNVTCEGCKGFFRRAMRNALRCPPR-KGACETTRKTRROCC 95
Db 74 GHELCRVCGDASGPHYVNLSCGCKGFFRRSVHGGAGRYACRGSGTCOMDAFMRKCO 133

Tue Feb 20 09:10:18 2001

us-09-143-828-2.rai

Page 10

[illegible]

Search completed: February 18, 2001, 14:31:58
Job time: 6201 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 18, 2001, 12:06:00 ; Search time 51.75 Seconds

(Without alignments)
270.834 Million cell updates/sec

Title: US-09-143-828-2
Perfect score: 2290
Sequence: 1 MEVPRKESNMHADFVHCEDT.....QDIHPATPLMDELFGITGS 434

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 88757 seqs, 32294092 residues

Total number of hits satisfying chosen parameters: 88757

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database: SwissProt_39

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived from analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	2390	100.0	434	1	PXR_HUMAN
2	1779.5	77.7	431	1	PXR_MOUSE
3	1761.5	76	431	1	PXR_PAT
4	815.5	35.6	448	1	VDR_COTVA
5	815	35.6	451	1	VDR_CHICK
6	814	35.5	422	1	VDR_XENLA
7	801	35.0	423	1	VDR_RAT
8	792.5	34.6	422	1	VDR_MOUSE
9	789	34.5	424	1	VDR_BOVIN
10	789	34.5	427	1	VDR_HUMAN
11	768	33.5	448	1	NR13_HUMAN
12	729	31.8	358	1	NR13_RAT
13	725	31.7	358	1	NR13_MOUSE
14	498.5	21.6	445	1	NRH3_MOUSE
15	498	21.7	461	1	NRH2_HUMAN
16	496.5	21.7	445	1	NRH3_RAT
17	491.5	21.5	447	1	NRH3_HUMAN
18	489	21.4	446	1	NRH2_MOUSE
19	487	21.3	446	1	NRH2_RAT
20	455.5	19.9	395	1	THB_PAROL
21	450	19.7	369	1	THB_CHICK
22	450	19.7	461	1	THB1_HUMAN
23	449	19.6	373	1	THB_BANCA
24	448	19.6	373	1	THB1_XENLA
25	445	19.4	476	1	THB2_HUMAN
26	444.5	19.4	757	1	ERL_DCCU
27	443	19.3	414	1	THB1_XENLA
28	442	19.3	675	1	ECR_AEDAE
29	441	19.3	461	1	THB1_MOUSE
30	441	19.3	475	1	THB2_MOUSE
31	437	19.1	461	1	THB1_RAT
32	437	19.1	514	1	THB2_RAT
33	434	19.0	416	1	THA_PAROL

34	433.5	18.9	427	1	THA1_BRARE	O98867 brachydanio
35	432	18.9	878	1	ECR_DROME	P34021 drosophila
36	428	18.7	448	1	RRE2_HUMAN	P10826 homo sapien
37	427.5	18.7	458	1	RRA_XENLA	P51126 xenopus lae
38	427.5	18.7	580	1	RORC_HUMAN	P51449 homo sapien
39	426.5	18.6	556	1	ECR_MANSE	P49883 manduca sex
40	425.5	18.6	408	1	THA_CHICK	P04625 gallus gall
41	425	18.6	455	1	RRE_CHICK	P22448 gallus gall
42	423.5	18.5	402	1	THA_APPRA	O42295 aptenodytes
43	423.5	18.5	402	1	THA_PYGAD	O42450 pygocelis
44	421.5	18.4	458	1	RRA_MOTVI	P18514 notophthalm
45	421.5	18.4	536	1	ECR_CHITE	P49882 chironomus

ALIGNMENTS

RESULT	ID	PXR_HUMAN	STANDARD	PRT	434 AA.
AC	075469				
DT	15-JUL-1999	(Rel. 38, Created)			
DT	15-JUL-1999	(Rel. 38, Last sequence update)			
DT	30-MAY-2000	(Rel. 39, Last annotation update)			
DE	ORPHAN NUCLEAR RECEPTOR PXR (PREGNANE X RECEPTOR) (ORPHAN NUCLEAR RECEPTOR PARL).				
GN	NR112 OR PXR.				
OS	Homo sapiens (human).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=LIVER;				
RC	MEDLINE=98395173; PubMed=9727070;				
RA	Lehmann J.M., Kckee D.D., Watson M.A., Willson T.M., Moore J.T.,				
RA	Kliwer S.A.;				
RT	"The human orphan nuclear receptor PXR is activated by compounds that regulates CYP3A4 gene expression and cause drug interactions.";				
RL	J. Clin. Invest. 102:1016-1023(1998)				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=LIVER;				
RC	MEDLINE=98445350; PubMed=9770465;				
RA	Bertilsson G., Heidrich J., Svensson K., Asman M., Jendeborg L.,				
RA	Sydney-Bachman M., Ohlsson R., Postlind H., Blomquist P.,				
RA	Berkestrand A.;				
RT	"Identification of a human nuclear receptor defines a new signalling pathway for CYP3A induction.";				
RT	Proc. Natl. Acad. Sci. U.S.A. 95:12208-12213(1998).				
CC	- FUNCTION: ORPHAN RECEPTOR: ITS NATURAL LIGAND IS PROBABLY PREGNANE. BINDS TO A RESPONSE ELEMENT IN THE CYP3A4 GENE PROMOTER.				
CC	- SUBUNIT: FORMS A HETERODIMER WITH RXR.				
CC	- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).				
CC	- TISSUE SPECIFICITY: EXPRESSED IN LIVER, COLON, AND SMALL				
CC	INTESTINE.				
CC	- INDUCTION: ACTIVATED BY NATURALLY OCCURRING STEROIDS SUCH AS				
CC	PREGNENONE AND PROGESTERONE.				
CC	- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.				
CC	NRL SUBFAMILY.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/				
CC	or send an email to license@sib-sib.ch).				
CC	-----				
CC	EMBL: AF061056; AAD05436.1; -				
CC	EMBL: AF084645; AAC64558.1; -				
CC	MIM: 603065; -				
CC	INTERPRO: IPR000324; -				
CC	INTERPRO: IPR000536; -				

DR INTERPRO: IPR001628; -
 DR INTERPRO: IPR001723; -
 DR PFAM: PF00104; hormone_rec; 1.
 DR PFAM: PF00105; zf-C4; 1.
 DR PRINTS: PRO0047; STEROIDFINGER.
 DR PRINTS: PRO0350; VITAMINDR.
 DR PRINTS: PRO0398; STROHOMONER.
 DR PROSITE: PS00031; NUCLEAR_RECEPTOR; 1.
 DR Receptor; Transcription regulation; DNA-binding; Nuclear protein;
 KW zinc-finger.
 FT DNA_BIND 41 107 C4-TYPE ZINC FINGERS (TWO).
 FT ZN_FING 41 61 C4-TYPE.
 FT ZN_FING 77 102 C4-TYPE.
 FT DOMAIN 108 204 HINGE.
 FT DOMAIN 205 434 LIGAND-BINDING.
 SQ SEQUENCE 434 AA; 49761 MW; 1DF6A2AE3109CADA CRC64;

Query Match 100.0%; Score 2290; DB 1; Length 434;
 Best Local Similarity 100.0%; Pred. No. 1.3e-176; Indels 0; Gaps 0;
 Matches 434; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEVPRKESWNRADVFHCEDETSYVGRKPSVNADEVGPGQICRYCGDKATGYHFNVTCEG 60
 DB 1 MEVPRKESWNRADVFHCEDETSYVGRKPSVNADEVGPGQICRYCGDKATGYHFNVTCEG 60

QY 61 CKGFFRRAMKRNARLCPFRKACETTRKTRROCCRLKCLSGKKKMMIMSDAVEE 120
 DB 61 CKGFFRRAMKRNARLCPFRKACETTRKTRROCCRLKCLSGKKKMMIMSDAVEE 120

QY 121 RRAIRKRSKSRGTOTPLGVGGLTEBOGMMIRELMDQMKTPTTSHFNFLPGVLS 180
 DB 121 RRAIRKRSKSRGTOTPLGVGGLTEBOGMMIRELMDQMKTPTTSHFNFLPGVLS 180

QY 121 RRAIRKRSKSRGTOTPLGVGGLTEBOGMMIRELMDQMKTPTTSHFNFLPGVLS 180
 DB 121 RRAIRKRSKSRGTOTPLGVGGLTEBOGMMIRELMDQMKTPTTSHFNFLPGVLS 180

QY 181 GCELPESIQAPSRREAAKMSQVRDLCSLVSQILRGEDSVNWKPADSGKEIFSL 240
 DB 181 GCELPESIQAPSRREAAKMSQVRDLCSLVSQILRGEDSVNWKPADSGKEIFSL 240

QY 241 PHADMSIYMKGIISFAKVISYRDLPIEDQISLKGAFELCOLFNTVNAETGWE 300
 DB 241 PHADMSIYMKGIISFAKVISYRDLPIEDQISLKGAFELCOLFNTVNAETGWE 300

QY 301 CGRLSYCEDTAGGFQOLLEPMLKFMKLKQLHHEEYVLMQALSLFSPDRGVLOHR 360
 DB 301 CGRLSYCEDTAGGFQOLLEPMLKFMKLKQLHHEEYVLMQALSLFSPDRGVLOHR 360

QY 361 VVDQLQEQFATLTKSYTECNRPQAHNPLFLKIMAMTELKSIINAQHTORLLRIODIHPF 420
 DB 361 VVDQLQEQFATLTKSYTECNRPQAHNPLFLKIMAMTELKSIINAQHTORLLRIODIHPF 420

QY 421 ATPLMQELFGITGS 434
 DB 421 ATPLMQELFGITGS 434

RESULT 2
 PKR_MOUSE STANDARD; PRT; 431 AA.
 ID PKR_MOUSE
 AC 054915;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE ORPHAN NUCLEAR RECEPTOR PKR (PREGNANE X RECEPTOR).
 GN NR112 OR PKR.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 RN [1]
 RP SEQUENCE FROM N.A. AND ALTERNATIVE SPLICING.
 RC TISSUE-LIVER.
 RX MEDLINE-98149345; PubMed-9489701;
 RX Kliever S.A., Moore J.T., Wade T., Staudinger J.L., Watson M.A.,
 RA Jones S.A., McKee D.D., Oliver B.B., Willson T.M., Zetterstrom R.H.,

RA Perlmann T., Leimann J.M.;
 RT "An orphan nuclear receptor activated by pregnanes defines a novel
 RT steroid signaling pathway."
 RL Cell 92:73-82(1998).
 CC -1- FUNCTION: ORPHAN RECEPTOR; ITS NATURAL LIGAND IS PROBABLY
 CC PREGNANE. BINDS TO A RESPONSE ELEMENT IN CYP1A GENES PROMOTER.
 CC -1- SUBUNIT: FORMS A HETERODIMER WITH RXR.
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; 1/PKR.1 (SHOWN HERE) AND
 CC 2/PKR.2, ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- INDUCTION: ACTIVATED BY NATURALLY OCCURRING STEROIDS SUCH AS
 CC PREGNANOLONE AND PROGESTERONE, SYNTHETIC GLUCOCORTICOID AND
 CC ANTIGLUCOCORTICOID AND 16-ALPHA-CARBONITRILE (PCN).
 CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.
 CC NRI SUBFAMILY.

 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

CC EMBL: AF031814; AAC39964.1; -
 CC MGI: 1337040; NR112.
 DR INTERPRO: IPR000324; -
 DR INTERPRO: IPR000536; -
 DR INTERPRO: IPR001628; -
 DR INTERPRO: IPR001723; -
 DR PFAM: PF00104; hormone_rec; 1.
 DR PFAM: PF00105; zf-C4; 1.
 DR PRINTS: PRO0047; STEROIDFINGER.
 DR PRINTS: PRO0350; VITAMINDR.
 DR PRINTS: PRO0398; STROHOMONER.
 DR PROSITE: PS00031; NUCLEAR_RECEPTOR; 1.
 DR Receptor; Transcription regulation; DNA-binding; Nuclear protein;
 KW zinc-finger; Alternative splicing.
 FT DNA_BIND 38 104 C4-TYPE ZINC FINGERS (TWO).
 FT ZN_FING 38 58 C4-TYPE.
 FT ZN_FING 74 99 C4-TYPE.
 FT DOMAIN 105 201 HINGE.
 FT DOMAIN 202 431 LIGAND-BINDING.
 FT VARSPLIC 171 211 MISSING (IN ISOFORM 2).
 SQ SEQUENCE 431 AA; 49567 MW; F592AF91F689329E CRC64;

Query Match 77.7%; Score 1779.5; DB 1; Length 431;
 Best Local Similarity 77.2%; Pred. No. 1.3e-135; Indels 1; Gaps 1;
 Matches 332; Conservative 42; Mismatches 55; Indels 1; Gaps 1;

QY 3 VAPKESWNRADVFHCEDETSYVGRKPSVNADEVGPGQICRYCGDKATGYHFNVTCEG 62
 DB 1 VAPKESWNRADVFHCEDETSYVGRKPSVNADEVGPGQICRYCGDKATGYHFNVTCEG 62

QY 63 GFFRRAMKRNARLCPFRKACETTRKTRROCCRLKCLSGKKKMMIMSDAVEE 122
 DB 63 GFFRRAMKRNARLCPFRKACETTRKTRROCCRLKCLSGKKKMMIMSDAVEE 122

QY 123 ALIKRKRKSRGTOTPLGVGGLTEBOGMMIRELMDQMKTPTTSHFNFLPGVLSGC 182
 DB 123 ALIKRKRKSRGTOTPLGVGGLTEBOGMMIRELMDQMKTPTTSHFNFLPGVLSGC 182

QY 183 ELPELIQAPSRREAAKMSQVRDLCSLVSQILRGEDSVNWKPADSGKEIFSLPH 242
 DB 183 ELPELIQAPSRREAAKMSQVRDLCSLVSQILRGEDSVNWKPADSGKEIFSLPH 242

QY 243 MADMSIYMKGIISFAKVISYRDLPIEDQISLKGAFELCOLFNTVNAETGWE 302
 DB 243 MADMSIYMKGIISFAKVISYRDLPIEDQISLKGAFELCOLFNTVNAETGWE 302

QY 303 RUSYCEDTAGGFQOLLEPMLKFMKLKQLHHEEYVLMQALSLFSPDRGVLOHRV 362
 DB 303 RUSYCEDTAGGFQOLLEPMLKFMKLKQLHHEEYVLMQALSLFSPDRGVLOHRV 362

DR INTERPRO; IPR000536; -
 DR INTERPRO; IPR001628; -
 DR PFAM; PF00104; hormone_rec; 1.
 DR PFAM; PF00105; zf-C4; 1.
 DR PRINTS; PR00047; STEROIDFINGER.
 DR PRINTS; PR000350; VITAMINDR.
 DR PROSITE; PS00031; NUCLEAR_RECEPTOR; 1.
 DR PROSITE; Transcription regulation; DNA-binding; Nuclear protein;
 KW Zinc-finger; Alternative Initiation.
 FT CHAIN 1 448
 FT INIT_MET 26 448
 FT ZN_BIND 44 109
 FT ZN_FING 44 64
 FT ZN_FING 80 104
 FT DOMAIN 110 212
 FT DOMAIN 213 448
 FT SEQUENCE 448 AA; 50668 MW; FCF1FC3DEAEAF3E CRC64;

Query Match 35.6%; Score 815.5; DB 1; Length 448;
 Best Local Similarity 40.9%; Pred. No. 3.3e-58;

Matches 179; Conservative 72; Mismatches 154; Indels 33; Gaps 10;

QY 17 CPTPE---SYPRKPSVNADE-EVGGPQICRGSDKATGYHFNVMCEGCKGFRRAMKKN 72
 DB 16 CSEQLQSSDMETPAVGTPEDFNVPICGVCGRATGFHFNAETCEGCKGFRRAMKKN 75
 QY 73 ARRCPRKACETITRTKRCQACRLKCLSGMKKEMISDEAVEERRALIKRKSR 132
 DB 76 AMTCEP-SQDCKITDNRHCOACRLKRCVDIGMKEFLIDEEYGRREMLKKEE 134
 QY 133 TGTPLGVQGLTEGQRMRELDAQMTFTDTSFKNFRLP-----GTL 178
 DB 135 ALKESLKP-LSEEOQVNIILFAHKKTFDITYSDFNFRPRVSKSSSTATSSSV 193
 QY 179 SSGCELPEISQAQPSREAAWQVRKDLCKLVSQLRGDSVWYKPPADSGKREIS 238
 DB 194 SQPSSSDNDVFGSDAFAPPEPOMFSNDLSESDSESSNMILPH-----LP 246
 QY 239 ILPHMADMTYMERKGISPAKVISYFDDLPIDEDQISLKGAFELCOLRNTVNAETGT 298
 DB 247 MFLHLDLVYSIQKVIKFAKMPGFRDLAEDQIALKSSALEVIMLSNOSTMEDMS 306
 QY 299 WEGCR--LSTCLSD-TAGGQQLLPRMFKFHMKRLDHEEYVIMQALISFSDRPG 355
 DB 307 WTCGSDNFKKYVDYQAGHSMLEPLVKEVGLKLNHEEHVLMALICLSPDRG 366
 QY 356 VLAHRVYDQLOEQFATLKSYIECNRPQAFRLFLTKIMAMTELNSINAOHTQ--RLR 413
 DB 367 VQDTSLVESIQDRSLDTLQYIINCRHPPPSRLIYAMIKLADLSINSHSKYRCIS 426
 QY 414 IODIHPF-ATPLMOELFG 430
 DB 427 FOPEHSMQTLPLVLEVG 444

RESULT 5
 VDR_CHICK STANDARD; PRT; 451 AA.
 AC 042392;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE VITAMIN D3 RECEPTOR (VDR) (1.25-DIHYDROXYVITAMIN D3 RECEPTOR).
 GN VDR OR N111.
 OS Gallus gallus (Chicken).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 CC Gallus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-LEGHORN; TISSUE-KIDNEY;

RX MEDLINE-97223369; PubMed-9056239;
 RA Lu Z., Hanson K., Deluca H.F.;
 RT Cloning and origin of the two forms of chicken vitamin D receptor.;
 RL Arch. Biochem. Biophys. 339:99-106(1997).

[2]
 RP SEQUENCE OF 45-114 FROM N.A.

RX MEDLINE-87149040; PubMed-3029866;
 RA McDonnell D.P., Mangelsdorf D.J., Pike J.W., Haussler M.R.,
 RA O'Malley B.W.;

RT Molecular cloning of complementary DNA encoding the avian receptor
 for vitamin D.

RL Science 235:1214-1217(1987).
 CC -1- FUNCTION: VDR MEDIATES THE ACTION OF VITAMIN D3 BY CONTROLLING
 THE EXPRESSION OF HORMONE SENSITIVE GENES.

CC -1- SUBCELLULAR LOCATION: NUCLEAR.

CC -1- ALTERNATIVE PRODUCTS: TWO FORMS (A AND B) ARE PRODUCED BY USE OF
 ALTERNATIVE INITIATION CODONS.

CC -1- TISSUE SPECIFICITY: EXPRESSED IN KIDNEY AND INTESTINE.

CC -1- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
 A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.

CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.

CC N1 SUBRAMITY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation
 at the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

CC EMBL; AF011356; AAB62579.1; -

DR INTERPRO; IPR000324; -

DR INTERPRO; IPR000536; -

DR INTERPRO; IPR001628; -

DR PFAM; PF00104; hormone_rec; 1.

DR PFAM; PF00105; zf-C4; 1.

DR PRINTS; PR00047; STEROIDFINGER.

DR PRINTS; PR000350; VITAMINDR.

DR PROSITE; PS00031; NUCLEAR_RECEPTOR; 1.

DR Receptor; Transcription regulation; DNA-binding; Nuclear protein;

KW Zinc-finger.

FT CHAIN 1 451

FT INIT_MET 15 451

FT ZN_BIND 47 112

FT ZN_FING 47 67

FT ZN_FING 83 107

FT DOMAIN 113 215

FT SEQUENCE 216 451

QY 8 SWNH-----ADFHCEPTESVPGKPSVNADEVGPGQICRGSDKATGYHFNVT 57

DB 7 SMDEOQSMALYPRDMDTVAASTSLP-DPAGDFGRNY--PRICGVCGRATGHNAMT 63

QY 58 CEGCKGFFRRAMKRNARLRCPPKACETITRTKRCQACRLKCLSGMKKEMISDEA 117

DB 64 CEGCKGFFRRAMKRNARLRCPPKACETITRTKRCQACRLKCLSGMKKEMISDEE 122

QY 118 VERBRLIKRKSEKRTGTOPLGVOGLTEGQRMRELDAQMTFTDTSFKNFRLP-- 175

DB 123 VQRKEMILKKEEALLESKLPK-LSEEOQVNIILFAHKKTFDITYSDFNFRPRVR 181

QY 176 -----GVLSSGCELPEISQAQPSREAAWQVRKDLCKLVSQLRGDSVW 223

DB 182 SKFSSMAITHSSSVYQSPSSSDNDVFGSDAFAPPEPOMFSNDLSESDSESSM 241

Query Match 35.6%; Score 815; DB 1; Length 451;
 Best Local Similarity 40.0%; Pred. No. 3.6e-58;
 Matches 181; Conservative 75; Mismatches 155; Indels 42; Gaps 11;

QY 224 NKPPADSGGKEITSLPHMADSTYMEKGIISFAKIVSYERDLPIDEDISLKGAFEL 283
 DB 242 NIEPH-----LPLMPLHADIVSISIOKVIKAKMIPGRDLTADQIALKSSAIEV 294
 QY 284 COLNENYVNAETGWEGR--LSTCLEL-TAGFQOLLPEPLKFKYMAKKTOLHEEY 340
 DB 295 IMLNSOFTWEDMSWTCGSNDKFKVSDVQACHSMDLEPLKFKYMAKKTOLHEEY 354
 QY 341 VLMOAISLSPDRPGVLOHVRVDOLOFOFATLTKSYIECNRPPOPHRFLFKTAMLT 400
 DB 355 VLMAICISLSPDRPGVDTSLVESIODRLSDIOTYIRCHNPPGSHLYAKMIQKLADL 414
 QY 401 RSTNAQHTQ--RLRLIDIHPE-ATPLMOELFG 430
 DB 415 RSLNEHSKORYCLSFQPSHMOJLPLVLEVF 447

RESULT 6
 VDR_XENLA STANDARD; PRT; 422 AA.
 ID VDR_XENLA 013124;
 AC 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE VITAMIN D3 RECEPTOR (VDR) (1,25-DIHYDROXYVITAMIN D3 RECEPTOR).
 GN VDR OR NR11.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipiloidea; Pipidae;
 OC Xenopodidae; Xenopus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-KIDNEY;
 RX MEDLINE-97307679; PubMed-9165021;
 RA Li Y.C., Bergwitz C., Jueppner H., Demay M.B.;
 RT Cloning and characterization of the vitamin D receptor from Xenopus
 laevis.
 RT Endocrinology 138:2347-2353(1997).
 RL -1- FUNCTION: VDR MEDIATES THE ACTION OF VITAMIN D3 BY CONTROLLING
 CC THE EXPRESSION OF HORMONE SENSITIVE GENES.
 CC -1- SUBCELLULAR LOCATION: NUCLEAR.
 CC -1- TISSUE SPECIFICITY: DETECTED IN ALL THE TISSUE EXAMINED. HIGHEST
 CC LEVEL IN SMALL INTESTINE AND SKIN
 CC -1- DEVELOPMENTAL STAGE: FIRST DETECTED AT STAGE 13. INCREASES
 CC GRADUALLY AND PEAKS AT STAGE 57-61 THEN DECREASES TO THE LEVEL
 CC SEEN IN ADULT.
 CC -1- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
 CC A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
 CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.
 CC NRI SUBFAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib.ch).
 CC
 CC EMBL: U91846; AAB58585.1;
 DR INTERPRO: IPR000324;
 DR INTERPRO: IPR000336;
 DR INTERPRO: IPR001628;
 DR PFAM: PF00104; hormone-rec; 1.
 DR PFAM: PF00105; zf-C4; 1.
 DR PRINTS: PRO0047; STROIDFINGER.
 DR PRINTS: PRO0350; VITAMINDR.
 DR PROSITE: PS00031; NUCLEAR RECEPTOR: 1.
 KW Receptor; Transcription regulation; DNA-binding; Nuclear protein;
 KW zinc-finger.
 FT DNA_BIND 25 90 C4-TYPE ZINC FINGERS (TWO).
 FT ZN_BIND 25 45 C4-TYPE.
 FT ZN_FING 61 85 C4-TYPE.

FT DOMAIN 91 188 HINGE.
 FT DOMAIN 188 422 LIGAND-BINDING.
 SQ SEQUENCE 422 AA: 48188 MW: C8A9F25414FEB9D5 CRC64;

Query Match 35.5% Score 814: DB 1: Length 422;
 Best local similarity 41.1%: Pred. No. 4e-58;
 Matches 172: Conservative 74: Mismatches 123: Indels 50: Gaps 10;

QY 38 POICRYCGDKATGYHFNWTCGCKGFFRRAMKRNARLRCEFRKGCACETRTKTRQCAC 97
 DB 22 PRICGCGDKATGYHFNWTCGCKGFFRRAMKRNARLRCEFRKGCACETRTKTRQCAC 80
 QY 98 RLKRLSEKMKEMIMSDENAEVERBALIKRKSEGTGTOPAGVGLTEOQRMIRELMDA 157
 DB 81 RLKRCVDIGMKKEFILTDEEVOKRQKMKRSEKALSKEMPR-1SDQOKMIDILDEA 139
 QY 158 QMKTFDTTSHKRNRLDGLVSSGCELPESLQAPREEMAKRSQVRKDCSLKYSIOLRG 217
 DB 140 HKRTFTTYSDEKRR-----PPYRENVDFRRITR-----SSVHTQG 178
 QY 218 ---EDGSWNKPPADS-----GKNE-IFSLPHMADSTYMEKGIIS 256
 DB 179 SPSEDSDFVTSPPDSEHGFASLFGQFEXSSMGKRGELSMPLHADIVSISIOKIG 238
 QY 257 FAKVSYFRDLPIEDISLKGAFELCOLFNVTENETWEGR--LSTCLEL-TAG 313
 DB 239 FAKMIPGRDLAEQDILLSSVIEYIMLRNSOFSLDDMSWTCGSEDFYKVDVTOA 298
 QY 314 GFOOLLEPKLFHYMLKKTOLHEEYVLMOAISLSPDRPGVLOHVRVDOLOFOFATL 373
 DB 299 GNNELLEPIKVFQYGLKKTOLHEEYVLMOAISLSPDRPGVLOHVRVDOLOFOFATL 358
 QY 374 KSYICNPPCAHREFLKTKAMTELRSINAQHTORLRLIDIHPPA---TPLMOELF 429
 DB 359 QYIICKHPPGSHLYAKMIQKLADLRLSNEHSKORYCLSFQPSHMOJLPLVLEVF 417

RESULT 7
 VDR_RAT STANDARD; PRT; 423 AA.
 ID VDR_RAT P13053;
 AC P13053;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-JAN-1990 (Rel. 13, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE VITAMIN D3 RECEPTOR (VDR) (1,25-DIHYDROXYVITAMIN D3 RECEPTOR).
 GN VDR OR NR11.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-89071726; PubMed-2849110;
 RA Burmester J.K., Wiese R.J., Maeda N., Deluca H.F.;
 RT "Structure and regulation of the rat 1,25-dihydroxyvitamin D3
 receptor."
 RT Proc. Natl. Acad. Sci. U.S.A. 85:9499-9502(1988).
 RL [2]
 RP SEQUENCE OF 58-423 FROM N.A.
 RX MEDLINE-88124963; PubMed-2829212;
 RA Burmester J.K., Maeda N., Deluca H.F.;
 RT "Isolation and expression of rat 1,25-dihydroxyvitamin D3 receptor
 cDNA."
 RT Proc. Natl. Acad. Sci. U.S.A. 85:1005-1009(1988).
 CC -1- FUNCTION: VDR MEDIATES THE ACTION OF VITAMIN D3 BY CONTROLLING
 CC THE EXPRESSION OF HORMONE SENSITIVE GENES.
 CC -1- SUBCELLULAR LOCATION: NUCLEAR.
 CC -1- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
 CC A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
 CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.
 CC NRI SUBFAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: J04147; AAA41089.1; -
 DR PIR: A31761; A31761.
 DR PIR: A31367; A31367.
 DR HSP: P03372; 1HCO.
 DR TRANSFAC: T00882; -
 DR INTERPRO: IPR000324; -
 DR INTERPRO: IPR000536; -
 DR INTERPRO: IPR001628; -
 DR PFAM: PF00104; hormone_rec; 1.
 DR PFAM: PF00105; zf-C4; 1.
 DR PRINTS: PR00047; STEROIDFINGER.
 DR PRINTS: PR00350; VITAMINDR.
 DR PROSITE: PS00031; NUCLEAR_RECEPTOR; 1.
 DR RECEPTOR: Transcription regulation; DNA-binding; Nuclear protein;
 KW Zinc-finger; Phosphorylation.
 FT DNA_BIND 24 89 C4-TYPE ZINC FINGERS (TWO).
 FT ZN_FING 24 89 C4-TYPE.
 FT ZN_FING 60 84 C4-TYPE.
 FT DOMAIN 90 187 HINGE.
 FT DOMAIN 188 423 LIGAND-BINDING.
 SQ SEQUENCE 423 AA; 47813 MW; 1A0B519A9DCCE990 CRC64;

Query Match 35.0%; Score 801; DB 1; Length 423;
 Best Local Similarity 42.0%; Pred. No. 4.5e-57;
 Matches 172; Conservative 75; Mismatches 135; Indels 28; Gaps 9;

QY 38 POICRYGDKATGYHFNWNTCEGCKGFFRRAMRNALRCPEFKGACITRTKTRCOAC 97
 DB 21 PRIGVCGDRATGFFHFNAMTCEGCKGFFRRSRKRAALTCPE-NGDCRITKDRRHCOAC 79
 QY 98 RLKRCLESGMKEMIMSDAVERERALKRKSEKSTGTOPLGVOGLTEORMMIRELMDA 157
 DB 80 RLKRCVIGIMKKEPILDEVOQRKREIMKREELALDLSRPK-LSSEQCHIIAIIILDA 138
 QY 158 QMKEFDTTFSEHKNRFLP-GVL-SSGCELPESIQAPSRERAAKWSQVRKDCISLKV- 211
 DB 139 HKHTYDPTVADRFPRPRVADSTGTSYSPR-PLTSSGNSSSSSSLYTTSLDME 194
 QY 212 -----SLQKRGEDGSVMNKPADSGKREITSLPHAMDMSTYFKGILISPAVISTFRD 266
 DB 195 PSFGSNLDLNGEDSD-----DPSVTLDLSPLSMLPHLADIVSYISIOKVIQGFAMKIPGRD 249
 QY 267 LPIDQISILKGAAPFELCOLRENTVFNNAETGTEGRLSYCLEDT---AGGFQQLLEPM 323
 DB 250 LITSDQIVLKSSAIEVIMLRSNOSTYMDMSWDCSGSDYIVDTVDYVSKAGHTLELEPL 309
 QY 324 LKFRMLKLIQLEHEEYVLMQALISLSPDRPGVLOHRYVDQLOEOPALITLKYTECNRPQ 383
 DB 310 IKFYGLKLLMLHEEHVILMAICIVSPDRGVODAKLVEAIODRLSTLQTYTRCHRP 369
 QY 384 PAHRLFLKTMATLSTASINAGTQ-RLLRIDIHFF-ATPLMOLFLG 430
 DB 370 PSHQVYAKMTOKLADLRLSLNEHRSKQYRSLSPENSMKLTPLVLEVF 419

RESULT 8
 VDR_MOUSE STANDARD; PRT: 422 AA.
 AC P48281;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE VITAMIN D3 RECEPTOR (VDR) (1,25-DIHYDROXYVITAMIN D3 RECEPTOR).
 GN VDR OR NR1H1.
 OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95137405; PubMed=7835717;
 RA Kamel Y., Kawada T., Fukuwatari T., Ono T., Kato S., Sugimoto E.,
 RT "Cloning and sequencing of the gene encoding the mouse vitamin D
 RT receptor."
 RL Gene 152:281-282(1995).
 CC -1- FUNCTION: VDR MEDIATES THE ACTION OF VITAMIN D3 BY CONTROLLING
 CC THE EXPRESSION OF HORMONE SENSITIVE GENES.
 CC -1- SUBCELLULAR LOCATION: NUCLEAR.
 CC -1- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
 CC A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
 CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.
 CC NRI SUBFAMILY.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: D31969; BAA06737.1; -
 DR HSP: P03372; 1HCO.
 DR MGD: MGI:103076; VDR.
 DR INTERPRO: IPR000324; -
 DR INTERPRO: IPR000536; -
 DR INTERPRO: IPR001628; -
 DR PFAM: PF00104; hormone_rec; 1.
 DR PFAM: PF00105; zf-C4; 1.
 DR PRINTS: PR00047; STEROIDFINGER.
 DR PRINTS: PR00350; VITAMINDR.
 DR PROSITE: PS00031; NUCLEAR_RECEPTOR; 1.
 DR RECEPTOR: Transcription regulation; DNA-binding; Nuclear protein;
 KW Zinc-finger; Phosphorylation.
 FT DNA_BIND 24 89 C4-TYPE ZINC FINGERS (TWO).
 FT ZN_FING 24 89 C4-TYPE.
 FT ZN_FING 60 84 C4-TYPE.
 FT DOMAIN 90 187 HINGE.
 FT DOMAIN 187 422 LIGAND-BINDING.
 SQ SEQUENCE 422 AA; 47851 MW; 4704CC8172445732 CRC64;

Query Match 34.6%; Score 792.5; DB 1; Length 422;
 Best Local Similarity 41.8%; Pred. No. 2.2e-56;
 Matches 174; Conservative 74; Mismatches 127; Indels 41; Gaps 11;

QY 38 POICRYGDKATGYHFNWNTCEGCKGFFRRAMRNALRCPEFKGACITRTKTRCOAC 97
 DB 21 PRIGVCGDRATGFFHFNAMTCEGCKGFFRRSRKRAALTCPE-NGDCRITKDRRHCOAC 79
 QY 98 RLKRCLESGMKEMIMSDAVERERALKRKSEKSTGTOPLGVOGLTEORMMIRELMDA 157
 DB 80 RLKRCVIGIMKKEPILDEVOQRKREIMKREELALDLSRPK-LSSEQCHIIAIIILDA 138
 QY 158 QMKEFDTTFSEHKNRFLP-GVLSSGCELP-----ESLQAPSR-EAAKWS 200
 DB 139 HKHTYDPTVADRFPRPRVADSTGTSYSPRPLTSSGNSSSSSSLYTTSLDMEPASFS 198
 QY 201 QVRKDCISLAKSLQKRGEDGSVMNKPADSGKREITSLPHAMDMSTYFKGILISPAVISTFRD 260
 DB 199 TM--DL-----NEGGS-----DDPSVTLDLSPLSMLPHLADIVSYISIOKVIQGFAMK 242
 QY 261 ISYFRDLPTEDQISILKGAAPFELCOLRENTVFNNAETGTEGRLSYCLEDT---AGGFQQL 317
 DB 243 IPEFRDLTSDQIVLKSSAIEVIMLRSNOSTYMDMSWDCSGSDYIVDTVDYVSKAGHTLELEPL 302
 QY 318 LKFRMLKLIQLEHEEYVLMQALISLSPDRPGVLOHRYVDQLOEOPALITLKYTECNRPQ 377
 DB 303 ELLEPLIKFYGLKLLMLHEEHVILMAICIVSPDRGVODAKLVEAIODRLSTLQTYTRCHRP 362

OY 378 ECNRPOHARFLFKIMAMTELRSINAOHTQ--RLLRIDIHPP-APLQMOELFG 430
 DB 363 RCHHPPGSHQYAKKIOKLADRLSLNEHSKOYRSLSPQENSKMLPVLVETVG 418

RESULT 9
 VDR_BOVIN STANDARD; PRT; 424 AA.

ID VDR_BOVIN STANDARD; PRT; 424 AA.
 AC 028037;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE VITAMIN D3 RECEPTOR (VDR) (1,25-DIHYDROXYVITAMIN D3 RECEPTOR).
 GN VDR OR NR1L1.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 RN (1)
 RP SEQUENCE FROM N.A.
 RX MEDLINE-97034797; PubMed-8880453;
 RA Nelberger H.L., Bosworth B.T., Reinhardt T.A.;
 RT "Nucleotide sequence of the bovine vitamin D3 receptor."
 RL J. Dairy Sci. 79:1313-1315(1996).
 CC -1- FUNCTION: VDR MEDIATES THE ACTION OF VITAMIN D3 BY CONTROLLING
 THE EXPRESSION OF HORMONE SENSITIVE GENES.
 CC -1- SUBCELLULAR LOCATION: NUCLEAR.
 CC -1- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
 A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
 CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.
 CC -1- NRI SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-slb.ch/announce/>
 CC or send an email to license@isb-slb.ch).
 CC -----
 CC EMBL; 050200; AAB01543.1; -.
 DR HSSP; P03372; IHCO.
 DR INTERPRO; IPR000324; -.
 DR INTERPRO; IPR000536; -.
 DR INTERPRO; IPR001628; -.
 DR PFAM; PF00104; hormone_rec; 1.
 DR PFAM; PF00105; zf-C4; 1.
 DR PRINTS; PRO0047; STROIDFINGER.
 DR PRINTS; PRO0350; VITAMINDR.
 DR PROSITE; PS00031; NUCLEAR_RECEPTOR; 1.
 DR Receptor; Transcription regulation; DNA-binding; Nuclear protein;
 KM Zinc-finger; Phosphorylation.
 FT ZINC_BIND 21 86
 FT ZN_FING 21 86
 FT ZN_FING 57 81
 FT ZN_FING 87 188
 FT DOMAIN 189 424
 FT SEQUENCE 424 AA; 47957 MW; E9E24526CE38C67D CRC64;

Query Match 34.5%; Score 789; DB 1; Length 424;
 Best Local Similarity 42.2%; Pred. No. 4,1e-56;
 Matches 174; Conservative 67; Mismatches 143; Indels 28; Gaps 9;

OY 38 POICRVCGRATGYHFNVMTCGCKGFRPARRMARLRCPFKAGCEITRTROQAC 97
 DB 18 PRICGCGDRATGFHFNAATCGCKGFRPARRMARLRCPFKAGCEITRTROQAC 76
 OY 98 RLKRCLESGKKKIMSDAVERRLIRKKKSERTGTPSLVQGLTEQRMKIRLMDA 157
 DB 77 RLKRCVIGMKKEFILDEVOQRKREMLIRKKEEALKOSLRPK-LSEQORITAILMDA 135

OY 158 QMKTFDTFFSHFNKFLRPLGVLSG----CELPSSLAQPSSEAKNSQVRKDC----- 207
 DB 136 HKHTYDETYEDFCQFRPVRVNDGCGSPSPRSNRTPT--FSGDSSSCSDHCITSSDM 193
 OY 208 ---SLKYSILCRGDSGVWYKPPADSGCKEIFSLPHMADMTYMKGLISPAKYSYF 264
 DB 194 MDSSFSNLTSEEDSD-----DPSTVLELSQJSMPLHDLVYSIQKYGAKMLPGF 248
 OY 265 RDLPIEDQISLTKGAFELCOLFNTVFNAETGWECCGRSLY---CLIEDYAGFOQLLE 321
 DB 249 RDLTSEQIVYLKSSALEVIMLRNENESTYMDKSWTCGNDYKRVSDYKASHLEIE 308
 OY 322 PMLFHYMLKQLQHEEYVLMQALISLSPDRGVQLHRRVDQLQEQFALTLSYIECNR 361
 DB 309 PLIKFVGKRLNHEEHVLMALICVSPDRGVQDAALIEAIDRLSTLTQYICRH 368
 OY 382 POPARHFLFKIMAMTELRSINAOHTQ--RLLRIDIHPP-APLQMOELFG 430
 DB 369 PPSHLLYAKMIQKLADRLSLNEHSKOYRSLSPQENSKMLPVLVETVG 420

RESULT 10
 VDR_HUMAN STANDARD; PRT; 427 AA.

ID VDR_HUMAN STANDARD; PRT; 427 AA.
 AC P11473;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-OCT-1989 (Rel. 12, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE VITAMIN D3 RECEPTOR (VDR) (1,25-DIHYDROXYVITAMIN D3 RECEPTOR).
 GN VDR OR NR1L1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-88217887; PubMed-2835767;
 RA Baker A.R., McDonnell D.P., Hughes M., Crisp T.M., Mangelsdorf D.J.,
 RA Haussler M.R., Pike J.W., Shine J., O'Malley B.W.;
 RT "Cloning and expression of full-length cDNA encoding human vitamin D
 RT receptor."
 RL Proc. Natl. Acad. Sci. U.S.A. 85:3294-3298(1988).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-92379083; PubMed-1324736;
 RA Goto H., Chen K.S., Prahl J.M., Deluca H.F.;
 RT "A single receptor identical with that from intestine/47D cells
 RT mediates the action of 1,25-dihydroxyvitamin D-3 in HL-60 cells."
 RL Blochim. Biophys. Acta 1132:103-108(1992).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX TISSUE-LENS EPITHELIUM;
 RA Rae J.L., Shepard A.R.;
 RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-97355582; PubMed-9212063;
 RA Miyamoto K., Kesterson R.A., Yamamoto H., Takekuni Y., Nishiwaki E.,
 RA Tatsumoto S., Inoue Y., Morita K., Takeda E., Pike J.W.;
 RT "Structural organization of the human vitamin D receptor chromosome
 RT gene and its promoter."
 RL Mol. Endocrinol. 11:1165-1179(1997).
 RN [5]
 RP SEQUENCE OF 24-90 FROM N.A.
 RX TISSUE-PERIPHERAL BLOOD; PubMed-1850412;
 RX MEDLINE-91210272; PubMed-1850412;
 RA Yu X.-P., Mochizuki H., Hustwyler F.G., Manolagas S.C.;
 RT "Vitamin D receptor expression in human lymphocytes. Signal
 RT requirements and characterization by western blots and DNA
 RT sequencing."
 RL J. Biol. Chem. 266:7588-7595(1991).
 RN [6]
 RP VARIANTS ASP-33; AND GLN-73.
 RX MEDLINE-89072761; PubMed-2849209;

FT VARIANT 50 50 R -> Q (IN RICKETS).
 FT VARIANT 73 73 /FTID-VAR_004661.
 FT /FTID-VAR_004662.

Query Match 34.5%; Score 789; DB 1; Length 427;
 Best Local Similarity 42.2%; Pred. No. 4.2e-56;
 Matches 174; Conservative 67; Mismatches 143; Indels 28; Gaps 9;

QY 38 POLCRVCGDKATGYHNVNMTCEGCKGFPRAMRNALACPRKGCETRTTRQOCAC 97
 DB 21 PLRCGCGGRRATGFHNNATCEGCKGFPRSKKRALFTCP-NGDCRTTKNRRHCQC 79
 QY 98 RLKRCLESGMKREKIMSDAVERRALIRKKSERTGTOPLGVGTEORRMIRELDA 157
 DB 80 RLKRCYDGMKKEFILTDEYQKRMILKREBEALKNLSPK-LSEQOQITILILDA 138
 QY 158 QMKTFTTSHKRNRLPGVLSG---CELPESIQAPSREAAKMSQVRKDL- 207
 DB 139 HHKTYDPTSDFCQFPPRVNDGGSHPSRPSNRTPS--PSGSSSSCSCHTSSDM 196
 QY 208 ---SLKVSILQAGEGVSNNYKPPADSGKEIFSLPHVADSTYMFGLISFAVISYF 264
 DB 197 NDSSESNLDLEEDSD-----DPVYTLSELSQLMPLADLVSTISIQVIGFAMKIGF 251
 QY 265 RDLPIEDQISLKGAFELCOLRFTVNAETGTEGRLSY---CLEDTAGGFOOLLE 321
 DB 252 RLUTSDQIVLLKSAIEVYIMLRSMESFTMDMSWTCGNODKYVSDVTKAGHLELE 311
 QY 322 PHLKFTYMKKQQLHEEYVLMQALSPDPKPGVLOHNVDOLOEORAITTKSTECNR 381
 DB 312 PIKRFQVQKRLNHEEYVLMALCIYSPDRGVODALILEADRLSNTQVTRGRH 371
 QY 382 POPARHFLKIMAMTELRSINAOHTO--RLRRIQ-DIHPRATPLMOELG 430
 DB 372 PPGSHLYAKMIQKLDRLSLNEHSKQYKCLSPQECMKLPLVLEVF 423

RESULT 11

NR13_HUMAN STANDARD; PRT; 348 AA.
 AC Q14994;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE ORPHAN NUCLEAR RECEPTOR NR113 (CONSTITUTIVE ANDROSTANE RECEPTOR) (CAR)
 DE (ORPHAN NUCLEAR RECEPTOR NB67).
 OS NR113.
 GN Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=LIVER;
 RA MEDLINE=94158627; PubMed=8114692;
 RA Baes M., Gullick T., Choi H.S., Martinoli M.G., Simha D., Moore D.D.,
 RT "A new orphan member of the nuclear hormone receptor superfamily that
 RT interacts with a subset of retinoic acid response elements.";
 RL MOL. CELL. BIOL. 14:1544-1552(1994).
 CC -1- FUNCTION: BINDS AND TRANSACTIVATES THE RETINOIC ACID RESPONSE
 CC ELEMENTS THAT CONTROL EXPRESSION OF THE RETINOIC ACID RECEPTOR
 CC BETA 2 AND ALCOHOL DEHYDROGENASE 3 GENES.
 CC -1- SUBUNIT: HETERODIMER OF NR113 AND RXR.
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
 CC -1- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN LIVER.
 CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.
 CC NR1 SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial

CC entitles requires a license agreement (see <http://www.isb-slb.ch/announce/>
 CC or send an email to license@isb-slb.ch).

DR EMBL: Z30425; CAA83016.1; -
 DR MIM: 603881; -
 DR INTERPRO: IPR000324; -
 DR INTERPRO: IPR000536; -
 DR INTERPRO: IPR001628; -
 DR INTERPRO: IPR001728; -
 DR PRAM: PR00104; hormone_rec; 1.
 DR PRAM: PR00105; zf-C4; 1.
 DR PRINTS: PR00047; STROIDFINGER.
 DR PRINTS: PR0035C; VITAMINDR.
 DR PRINTS: PR00546; THYROIDHOMR.
 DR PROSITE: PS00031; NUCLEAR RECEPTOR; 1.
 KW Receptor; Transcription regulation; DNA-binding; Nuclear protein;
 FT Zinc-finger; Activator.
 FT DNA_BIND 11 76 C4-TYPE ZINC FINGERS (TWO).
 FT ZN_FING 11 31 C4-TYPE.
 FT ZN_FING 47 71 C4-TYPE.
 SQ SEQUENCE 348 AA; 39558 MW; 92EAD07D7DCB9DA CRC64;

Query Match 33.5%; Score 768; DB 1; Length 348;
 Best Local Similarity 42.2%; Pred. No. 1.6e-54;
 Matches 164; Conservative 62; Mismatches 109; Indels 54; Gaps 6;

QY 41 CRVCGDKATGYHNVNMTCEGCKGFPRAMRNALACPRKGCETRTTRQOCACR 100
 DB 11 CVVCGDQATGYHNVNMTCEGCKGFPRSKKRALFTCP-AGSCVSKTORRCPACRQ 69
 QY 101 KCLSEGMKREKIMSDAVERRALIRKKSERTGTOPLGVGTEORRMIRELDAOK 160
 DB 70 KCLDAGKRMKMIISAEHLARAKQARAPQVQ-----LSKQDELIRLLAHNR 123
 QY 161 TDDTSHFRNRLPGVLSGCELPELSAPSRREAAKMSQVRKDLKLSVSLQAGEDG 220
 DB 124 HMGTFEQFVQFRPRLHFNHQ-PLPTLAP----- 153
 QY 221 SWNVKPPADSGKEIFSLPHVADSTYMFGLISFAVISYFRLPIEDQISLKGAA 280
 DB 154 -----VLPTHRADINTENVLOVIFETDLPAVRSPLPIEDQISLKGAA 198
 QY 281 FELCOLRFTVNAETGTEGRLSYCLEDTAG-GFQOOLLEPMLEFHKRLQLOHEE 339
 DB 199 VECHIVLNTTFCQONFCLPRLTIEDGAVGVELELFFHGTLRKLOQEP 258
 QY 340 YVLMQALSPDPKPGVLOHNVDOLOEORAITTKSYIECNRPAPARHFLKIMAMTE 399
 DB 259 YVLLAVALSPDRPGVTOHDEIDQLOEEMALTLQSYINGQQRPRDRFLYAKLGLAE 318
 QY 400 LRSINAOHTORLRIQDIHPRATPLMOEL 428
 DB 319 LRSINAVGYQIHOGLSAM-WPLDQEI 346

RESULT 12

NR13_RAT STANDARD; PRT; 358 AA.
 AC G9Q0S1;
 DT 01-OCT-2000 (Rel. 40, Created)
 DT 01-OCT-2000 (Rel. 40, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE ORPHAN NUCLEAR RECEPTOR NR113 (CONSTITUTIVE ANDROSTANE RECEPTOR)
 DE (CAR).
 GN NR113 OR CAR.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=WISTAR KYOTO, AND FISCHER; TISSUE=LIVER;
 RA Yoshinari K., Sueyoshi T., Moore R., Negishi M.;

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial


```

Db 11 ASEEYGRNVCYGDRTATGTHFHALTCGCGKFFRTVSKTIGPICPF-AGRCVSKAQ 69
QY 91 RROCOACRLRCLSGSKKMKKEMSDAEVRRRLIKRKSERGCTGPLVGITGEORAM 150
Db 70 RRRPACLOLCLVGRKMKITLAEALALRRARQORREKSLD-----LNDQKEL 123
QY 151 IREIMADOMKTEDTFSHFKNRFLPGVLSGCELPESLQAFSREAAKMSQVAKDCLSK 210
Db 124 VOILLGATHRHVGPLDOFVQFKPPAVL----- 151
QY 211 VSILREDSVWYKRRPADSGKEIFSLPFRADSTYKFGIISFANVISYFDLPIE 270
Db 152 -----FMHRRPQPRG--PVLPLTHPADINTFWOQIIFTKDLPLRSLTWE 198
QY 271 DOISLNGAAFEICQLRFVFNATGTCGRLSYCLSDTA-GGFQOILLRPMKLFHM 329
Db 199 DOISLNGAAVEIHLISLNTFLQIDENPFCGLCKMEDAVHAGQYEFELSLHFKN 258
QY 330 LKKQLHEEYVLMQATISLSPDRPGVLOHRYVDQLEQFAITLKSYLECNRPPAHRL 389
Db 259 LKGHLQEPYVLMATLSPDRPGVLOHRYVDQLEQFAITLKSYLECNRPPAHRL 318
QY 390 FLKIMAVLTELRSINQHRLRIODIHPRFPLMOEL 428
Db 319 YAKIMGLADLNINNAVSTELQRLSELSAM-TPLIGE1 356

RESULT 14
NRH3_MOUSE STANDARD: PRT: 445 AA.
AC 0920Y9:
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DE OXYSTEROLS RECEPTOR LXR-ALPHA (LAYER X RECEPTOR ALPHA) (NUCLEAR ORPHAN RECEPTOR LXR-ALPHA).
GN NRH3 OR LXR.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN (1)
RP SEQUENCE FROM N.A.
RA Chen Y.E., Horluchi M., Dzau V.J.;
RL Submitted (Aug-1998) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: ORPHAN RECEPTOR. INTERACTION OF LXR ALPHA WITH RXR SHIFTS RXR FROM ITS ROLE AS A SILENT DNA-BINDING PARTNER TO AN ACTIVE LIGAND-BINDING SUBUNIT IN MEDIATING RETINOID RESPONSES THROUGH TARGET GENES DEFINED BY LXRS (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: NUCLEAR (POTENTIAL).
CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.
CC -1- SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation- the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AF085745; AAD16050.1;
CC HSSP: P20393; 1A6Y.
CC INTERPRO: IPR000536;
CC INTERPRO: IPR001628;
CC PFM: PF00104; hormone_rec; 1.
CC PFM: PF00105; zf-C4; 1.
CC PRINTS: PR00047; STROIDFINGER.
CC PROSITE: PS00031; NUCLEAR_RECEPTOR; 1.
CC Zinc-finger.
CC Zinc-binding; DNA-binding; Nuclear protein;
CC DN_BIND 96 161 C4-TYPE ZINC FINGERS (TWO).
CC ZN_FING 96 116 C4-TYPE.

```

```

FT 2N_FING 132 156 C4-TYPE.
SQ SEQUENCE 445 AA; 50417 MW; 1A426DF38D935731 CRC64;

Query Match 21.8% Score 498.5; DB 1; Length 445;
Best Local Similarity 30.1%; Pred. No. 9,2e-33;
Matches 129; Conservative 85; Mismatches 143; Indels 71; Gaps 13;

QY 18 EDTEVPGKISVNADEEVGGPOICRYCGDKATGYHFNWATCGCGKGFPRBAKRVARLRC 77
Db 73 EPTELRQKRRKGPAPKMGNELSYCGDKAGFHYNVLSCGCGGFFRRSYIKARLYC 132
QY 78 PFRKGCETIRKRRQCCARLKLCSGKMKEMSDAEVRRRLIKRKS----- 131
Db 133 -HSGGCPMTYTKRRRCQCRRLKRCQAMRECVLSEQIHLK--LKQEEQQAQATS 189
QY 132 ---RTGTPGVGLTEEDRMMIRLMDOMKTEPDTFSHFKNRFLPGVLSGCELPESL 188
Db 190 VSPRVSSPCVLPQSLPBDLGMIEKLVAAQDCNRFS----- 228
QY 189 QAPSRBAARWQVRKDCSLRVSQLRGSDSVWYKRRPADSGKEIFSLPFRADST 248
Db 229 ---DRLRVTFW-PIAPD-----PQSRBAARQORFA---HFTELAI 260
QY 249 YMEKGIISPAKAYISYFRDLPIDQISILKGAAFELCOLRENTVNR--AETGTEGRLSY 306
Db 261 VSVQELVDPRKQPLPGFLQSLREDQIALKTSIEYMLLETERRRNPNGSESTTF-LKDFSY 319
QY 307 CLEDTA-GEFQDILLRPMKLFHYMLKRLQLEHEEYVLMQATISLSPDRPGVLOHRYVDOL 365
Db 320 NREDFAKAGLOVEFINPIEFERANNEQLNDAPALIAISISADPNVODQLOVERL 379
QY 366 QDQFAITLKSYLECNRPPAHRLFLKIMAVLTELRSINQHRL--LRDIHPRPATP 423
Db 380 QHTYVLAALVYSINHPDP--LMEPRMIAVLISLRTISVSHQVVALRLQD--KRLPP 435
QY 424 LMQELFGI 431
Db 436 LLSLEIMDV 443

RESULT 15
NRH2_HUMAN STANDARD: PRT: 461 AA.
AC P55055:
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DE OXYSTEROLS RECEPTOR LXR-BETA (LAYER X RECEPTOR BETA) (NUCLEAR ORPHAN RECEPTOR LXR-BETA) (UBIQUITOUSLY-EXPRESSED NUCLEAR RECEPTOR) (NUCLEAR RECEPTOR NER).
GN NRH2 OR LXR. OR UNR OR NER.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
RN (1)
RP SEQUENCE FROM N.A.
RA MEDLINE-95011628; Pubmed-7926814;
RA Shinar D.M., Endo N., Rutledge S.J., Vogel R., Rodan G.A., Schmidt A.;
RT "NER, a new member of the gene family encoding the human steroid hormone nuclear receptor."
RL Gene 147:273-276(1994).
CC -1- FUNCTION: ORPHAN RECEPTOR. BINDS PREFERENTIALLY TO DOUBLE-STRANDED OLIGONUCLEOTIDE DIRECT REPEATS HAVING THE CONSENSUS HALF-SITE SEQUENCE 5'-AGGTCA-3' AND 4-NT SPACING (DR-4).
CC -1- SUBUNIT: FORMS A HETERODIMER WITH RXR.
CC -1- SUBCELLULAR LOCATION: NUCLEAR (POTENTIAL).
CC -1- TISSUE SPECIFICITY: UBIQUITOUS.
CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.
CC -1- SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```


GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 18, 2001, 13:03:51 ; Search time 57.29 Seconds

(Without alignments)
514,381 Million cell updates/sec

Title: US-09-143-828-2

Perfect score: 2290

Sequence: 1 MEVRPKESWNHADVFCEDT.....ODHPFATPLMQLFGITGS 434

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 195891 seqs, 67900655 residues

Total number of hits satisfying chosen parameters: 195891

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database: PIR 66: 1: PIR1:*

2: PIR2:*

3: PIR3:*

4: PIR4:*

SUMMARIES

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Query Match Length	ID	Description
1	991.5	43.3	386 2	thyroid hormone re
2	817	35.7	420 2	vitamin D receptor
3	815.5	35.6	448 2	vitamin D receptor
4	801	35.0	423 2	1,25-dihydroxyvita
5	792.5	34.6	422 2	vitamin D receptor
6	790.5	34.5	425 2	vitamin D receptor
7	789	34.5	427 2	vitamin D receptor
8	789	34.5	427 2	vitamin D receptor
9	768	33.5	448 2	vitamin D receptor
10	498	21.7	461 2	steroid hormone re
11	496.5	21.7	445 2	steroid hormone re
12	491.5	21.5	447 2	steroid hormone re
13	489	21.4	446 2	steroid hormone re
14	487	21.3	446 2	steroid hormone re
15	456	19.9	446 2	steroid hormone re
16	451.5	19.7	484 2	steroid hormone re
17	450	19.7	369 1	steroid hormone re
18	450	19.7	369 1	steroid hormone re
19	449	19.6	373 2	steroid hormone re
20	448	19.6	373 2	steroid hormone re
21	445	19.4	456 1	steroid hormone re
22	443	19.3	373 2	steroid hormone re
23	441	19.3	461 2	steroid hormone re
24	441	19.3	461 2	steroid hormone re
25	441	19.3	475 2	steroid hormone re
26	435	19.0	464 2	steroid hormone re
27	434	19.0	416 2	steroid hormone re
28	434	19.0	448 2	steroid hormone re
29	432	18.9	878 2	steroid hormone re

Query Match	Score	Query Match Length	ID	Description
18	EDTSVQKRSVNADEYVGPQICRVCGDAGTGHVNTCEGCKGFFRRAMRNALRC	77		
14	EEEDASNSCGTGEDDDGDPICRACGDRATGYHFNAAMCEGKGFRRVAVRNRLSC	73		
78	PRKACETITRKTROCOACRLKRCLESGMKKREKIMSDAVERRALIKRK-KSERGTQ	136		
74	PF-QNSCVINKSNRHOACRLKCDIGKRLKIMSDAVERRALIKRNHLTKLPPT	132		
137	PLVQGLTEQRMKRLMDQMTEDTPTSHKRNFLPGVSSGCELPSLQAPSREEA	196		
133	PGGA-SLTPGQHLTQVGAKHTKTFEDNFTFSKNER-----PIR---	171		
197	AKMQVAKDCLSVSLQGLGEGSVNTPPADSGKKEFSLPLPHADNSTMEFGITS	256		
172	-----SSDPT---QEPQATS---SEAFMLPHISDLYVMIRGITS	207		
257	FAKVISYFDLPEDQISLKGAFELCOLRFVTVNAEGTWEGCRSLVSCLEDT-AGG	315		
208	FAKLPKFSKLDIEDQALALKGVAEVAIRFTVNSDNTWECGPFYTDIDMFLAGE	267		
316	QQLLEMLKFTYMLKLLQHEEYVLMQALISFSDRPGVLOHRYVDLOQOFATLKS	375		
268	ROLFLEPLVIRHMRKRLNLOSEYVAMMALSTFASDRPGVCKWEXIKLOEHLATLMD	327		
376	YIECNR-POPAPHFLFKIYAMLTSLRSTNAOCTOLLRODHPATLMOELG	430		
328	FIDSQRPSPQNLILYFKIMECLTELVNDIHSKOLLEIWDIOPATLMEVFG	383		

RESULT 2

Query Match 35.7%; Score 817; DB 2; Length 420;
 Best Local Similarity 43.1%; Pred. No. 3e-56; Indels 22; Gaps 12;
 Matches 179; Conservative 70; Mismatches 144;

Vitamin D receptor subtype a - Parathenchys olivaceus
 C:Species: Parathenchys olivaceus
 C:Date: 09-Jun-2000 #sequence_revision 09-Jun-2000 #text_change 18-Aug-2000
 C:Accession: J07229
 R:Suzuki, T.; Suzuki, N.; Sriwastava, A.S.; Kurokawa, T.
 Biochem. Biophys. Res. Commun. 270, 40-45, 2000
 A:Title: Identification of cDNAs encoding two subtypes of vitamin D receptor in flounder
 A:Reference number: J07229
 A:Accession: J07229
 A:Molecule type: mRNA
 A:Residues: 1-420 <SU>
 A:Cross-references: DDBJ:AB037673
 A:Experimental source: Intestine
 C:Comment: This receptor is an important factor in calcium homeostasis and bone formation
 C:Gene: Vdr-a
 C:Superfamily: unassigned erba-related proteins; erba transforming protein homology
 C:Keywords: bone; calcium transport; DNA binding; hormone receptor; intestine; vitamin D

Query Match 35.7%; Score 817; DB 2; Length 420;
 Best Local Similarity 43.1%; Pred. No. 3e-56; Indels 22; Gaps 12;
 Matches 179; Conservative 70; Mismatches 144;

Vitamin D receptor subtype a - Parathenchys olivaceus
 C:Species: Parathenchys olivaceus
 C:Date: 09-Jun-2000 #sequence_revision 09-Jun-2000 #text_change 18-Aug-2000
 C:Accession: J07229
 R:Suzuki, T.; Suzuki, N.; Sriwastava, A.S.; Kurokawa, T.
 Biochem. Biophys. Res. Commun. 270, 40-45, 2000
 A:Title: Identification of cDNAs encoding two subtypes of vitamin D receptor in flounder
 A:Reference number: J07229
 A:Accession: J07229
 A:Molecule type: mRNA
 A:Residues: 1-420 <SU>
 A:Cross-references: DDBJ:AB037673
 A:Experimental source: Intestine
 C:Comment: This receptor is an important factor in calcium homeostasis and bone formation
 C:Gene: Vdr-a
 C:Superfamily: unassigned erba-related proteins; erba transforming protein homology
 C:Keywords: bone; calcium transport; DNA binding; hormone receptor; intestine; vitamin D

Query Match 35.7%; Score 817; DB 2; Length 420;
 Best Local Similarity 43.1%; Pred. No. 3e-56; Indels 22; Gaps 12;
 Matches 179; Conservative 70; Mismatches 144;

Vitamin D receptor subtype a - Parathenchys olivaceus
 C:Species: Parathenchys olivaceus
 C:Date: 09-Jun-2000 #sequence_revision 09-Jun-2000 #text_change 18-Aug-2000
 C:Accession: J07229
 R:Suzuki, T.; Suzuki, N.; Sriwastava, A.S.; Kurokawa, T.
 Biochem. Biophys. Res. Commun. 270, 40-45, 2000
 A:Title: Identification of cDNAs encoding two subtypes of vitamin D receptor in flounder
 A:Reference number: J07229
 A:Accession: J07229
 A:Molecule type: mRNA
 A:Residues: 1-420 <SU>
 A:Cross-references: DDBJ:AB037673
 A:Experimental source: Intestine
 C:Comment: This receptor is an important factor in calcium homeostasis and bone formation
 C:Gene: Vdr-a
 C:Superfamily: unassigned erba-related proteins; erba transforming protein homology
 C:Keywords: bone; calcium transport; DNA binding; hormone receptor; intestine; vitamin D

F:42-362/Domain: erba transforming protein homology <ERBA>

Query Match 35.6%; Score 815.5; DB 2; Length 448;
 Best Local Similarity 40.9%; Pred. No. 4.2e-56;
 Matches 179; Conservative 72; Mismatches 154; Indels 33; Gaps 10;

17 CEDDE---SVGRKPSVNADE-ENGPGQICRGDRAFGYHFNWTCGCKGFFRRAMRN 72
 16 CESELOSSDDETFVAVGTPEFDRNVPICGVDRAFGHFNATCGCKGFFRRAMRN 75
 73 ARLCPFRKGAACEITRTKTRCOACRLKCLSGMKKEMTMSDAVEERRALIKRKSR 132
 76 AMFCPE-SGDCIKTKDNRHCQACRLKRCVDIGMKKEFILLDEYQRRKREMLIKREEE 134
 133 TGTPLQVQGTTEQGRMIRLMDAOKTDTTSHKKNRLP-----GVL 178
 135 ALKEELKPK-SEBQOKVITILLEAHHKTFDTYSDNKKRPFRKFSSTATSHSSV 193
 179 SSGCELPSILAPSRREAAKMSQVRKDLCSLKYSLQLRGDSVWNYKPPADSGKEIFS 238
 194 SODPSSESDNDFGSDAFGARFPEMEQMSNDLSESDSPSMNTELP-----LP 246
 239 ILPHNADSTTFNGIISFARVISYFDLPLEDOISILKGAAPFCOLRRNTVFNATGT 298
 247 MFLHDLVSYISQKVIQAFKMIPEFDLPLEDOIALKSSALEIVIMLRNQSFTMEDMS 306
 299 WECGR-LSYCLIED-TAGFPQOILEPMLKRYMLKLOLHEEYVLMQAIISLSPDRPG 355
 307 WTCGSDNRKRYVSDTQGHSMDLLEPLKRVQVGLKNTLHEEHVILMAICILSPDRPG 366
 356 VLQHRVVDLQEOFAITLTKSYIECNRPQAPARFLFKIMAMTELRSINAOHTO--RLR 413
 367 VQDTLVESIDRLSDTLQTIYIRCHHPRGSRLLYAKMIQIADLRSLNEHRSKQYRCLS 426
 414 IODIHPF-ATPLMDLEFG 430
 427 FQPEHSMQTLPLVLEVF 444

RESULT 4
 A:Residues: 1-423 <SU>
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 20-Sep-1999
 C:Accession: A31761; A31367
 R:Burnester, J.K.; Wiese, R.J.; Maeda, N.; Deluca, H.F.
 Proc. Natl. Acad. Sci. U.S.A. 85, 9499-9502, 1988
 A:Title: Structure and regulation of the rat 1,25-dihydroxyvitamin D-3 receptor.
 A:Reference number: A31761; MUID:89071726
 A:Accession: A31761
 A:Molecule type: mRNA
 A:Residues: 1-423 <SU>
 A:Cross-references: GB:J04147; GB:J03630; NID:g203956; PIDN:AAA41089.1; PID:g203957
 R:Burnester, J.K.; Maeda, N.; Deluca, H.F.
 Proc. Natl. Acad. Sci. U.S.A. 85, 1005-1009, 1988
 A:Title: Isolation and expression of rat 1,25-dihydroxyvitamin D-3 receptor cDNA.
 A:Reference number: A31367; MUID:88124963
 A:Accession: A31367
 A:Molecule type: mRNA
 A:Residues: 1-423 <SU>
 C:Superfamily: unassigned erba-related proteins; erba transforming protein homology
 C:Keywords: DNA binding; transcription regulation; zinc finger
 F:22-33/Domain: erba transforming protein homology <ERBA>
 F:24-44/Region: zinc finger
 F:60-84/Region: zinc finger

Query Match 35.0%; Score 801; DB 2; Length 423;
 Best Local Similarity 42.0%; Pred. No. 5.3e-55;
 Matches 172; Conservative 75; Mismatches 135; Indels 28; Gaps 9;

38 PQICRGVCGDKIIGYHFNWTCGCKGFFRRAMRNALRCPPKAGACETTRTRGQAC 97

Db 124 HMGMEFOVQFRPAHLFIHQ-PIETLAP----- 153
 Qy 221 SVMNKRPAVSGGKEIRSLPHMADSTYFKGIIISPAKVISYFRDLPIEDQISLKGAA 280
 Db 154 -----VLPETHRADITFQVNLVIRKFDLPFRSLPIEDQISLKGAA 198
 Qy 281 FELCQLNFTVFNATGTCWEGRLSTCYLEDTAG-GFOQLLEPMLKFKHYMLKRLQLEHEE 339
 Db 199 VEICHIYLVNTFELQTNFLCGLPLRTIEDGAVGFOVELLELFFHGTIRLQLOEPE 258
 Qy 340 YVLMQALISLSPRPVGLQHRVVDQLOEOPATLKSYIECNRPQAPRFLFKIMAMTE 399
 Db 259 YVLLAAALFSPRPVGLQHRVVDQLOEOPATLKSYIECNRPQAPRFLFKIMAMTE 318
 Qy 400 LRISNAHQRLRIODIHFPATPLMOEL 428
 Db 319 LRISNAHQRLRIODIHFPATPLMOEL 346

RESULT 10

JC4014
 steroid hormone-nuclear receptor NER - human
 C:Species: Homo sapiens (man)
 C:Date: 13-Jun-1995 #sequence_revision 14-Jul-1995 #text_change 24-Sep-1999
 C:Accession: JC4014
 R:Shinar, D.M.; Endo, N.; Rutledge, S.J.; Vogel, R.; Rodan, G.A.; Schmidt, A.
 Gene 147, 273-276, 1994
 A:Title: NER, a new member of the gene family encoding the human steroid hormone nuclear receptor
 A:Reference number: JC4014; MUID:95011628
 A:Accession: JC4014
 A:Molecule type: mRNA
 A:Residues: 1-461 <SH>
 A:Cross-references: GB:U07132; MID:9641961; PIDN:AAA61783.1; PID:9641962
 A:Experimental source: osteosarcoma cells SAOS-2/B10
 C:Genetics:
 A:Gene: GDB:UNR
 A:Cross-references: GDB:389570; OMIM:600380
 A:Map position: 19q13.3-19q13.3
 C:Superfamily: unassigned erba-related proteins; erba transforming protein homology
 C:Keywords: steroid hormone receptor
 F:85-381/Domain: erba transforming protein homology <ERBA>
 F:87-154/Domain: DNA binding #status predicted <BIN>

Query Match 21.7%; Score 498; DB 2; Length 461;
 Best Local Similarity 29.0%; Pred. No. 2.6e-31;
 Matches 126; Conservative 86; Mismatches 125; Indels 98; Gaps 14;

Qy 37 GPQIRVCGDKATGYHFNVTCEGCKGFFRRAMRNARLRCFPR-KGACETTRTRQCO 95
 Db 83 GHELCLRVCGDKATGYHFNVTCEGCKGFFRRAMRNARLRCFPR-KGACETTRTRQCO 142
 Qy 96 ACRLKCLSGKMKKIMSDAEVERRLIRKKSEGTGTO-PIGVG----- 142
 Db 143 QCRLLKCLSGKMKKIMSDAEVERRLIRKKSEGTGTO-PIGVG----- 142
 Qy 143 -----LTEORMMIRELMDAQMTFTDTSFHKFRPLPGVLSGCE 183
 Db 201 SPGSEAGSOGSGEGEVQLTMAQELMIQQLVAAQLQCNKRSFS----- 244
 Qy 184 LPESLQASREBAKWSQVRKDCSLKYSLOLRGEGSVWYKRPADSGGKEIFSLPHM 243
 Db 245 -----DQPVTFWP-----LGND-----PQSRDARQORFA---HF 271
 Qy 244 ADMSTYFKGIISPAKVISYFRDLPIEDQISLKGAEFLCQLRPNYFNATGTCWEC- 301
 Db 272 TELAIISVQELVDPAKQVPGFLQLOREDOILAKASTIEMLLETARRNHET---ECIT 328
 Qy 302 --GRISTYLED-TAGGFOQLLEPMLKFKHYMLKRLQLEHEEYVLMQALISLSPRPVGLQ 358
 Db 329 FLKDTYTSKDDFHRAGLQVEIFINPIEFSSRAMRRRLGLDDAEYALLAINIFSADRPNVQE 388

Qy 359 HRVVDQLOEOPATLKSYIECNRPQAPRFLFKIMAMTELRISNAHQRL--LRID 416
 Db 369 PERVALDQFVEALSTRIKRPQDOLR-FPRMIMKLVSLRTLSVSHSEVAFALRLOD 446
 Qy 417 IHFPATPLMOELFGI 431
 Db 447 --KKRPLSLSEIMDV 459

RESULT 11

A56043
 steroid hormone receptor-like protein RLD-1 - rat
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 01-Dec-1995 #sequence_revision 01-Dec-1995 #text_change 24-Sep-1999
 C:Accession: A56043
 R:Apfel, R.; Benbrook, D.; Lernhardt, E.; Ortiz, M.A.; Salbert, G.; Pehl, M.
 Mol. Cell. Biol. 14, 7025-7035, 1994
 A:Title: A novel orphan receptor specific for a subset of thyroid hormone-responsive
 A:Reference number: A56043; MUID:95021230
 A:Accession: A56043
 A:Status: Preliminary
 A:Molecule type: mRNA
 A:Residues: 1-445 <APF>
 A:Cross-references: GB:U11685; MID:9555751; PIDN:AAA3633.1; PID:9555752
 A:Note: authors translated the codon GAG for residue 73 as Ser
 C:Superfamily: unassigned erba-related proteins; erba transforming protein homology
 C:Keywords: steroid hormone receptor; zinc finger
 F:94-365/Domain: erba transforming protein homology <ERBA>

Query Match 21.7%; Score 496.5; DB 2; Length 445;
 Best Local Similarity 30.1%; Pred. No. 3.3e-31;
 Matches 129; Conservative 84; Mismatches 144; Indels 71; Gaps 13;

Qy 18 EDTESVGRKSYNADEEVGGPQICRVCGDKATGYHFNVTCEGCKGFFRRAMRNARLRC 77
 Db 73 EPTELRPQKRRKGPAPKMLGNELCSVCGDKASAFHYNVLSCEGCKGFFRRSVYKGRYIC 132
 Qy 78 PFRGACETTRTRQCOACRLKCLSGKMKKIMSDAEVERRLIRKKSE----- 131
 Db 133 -HSGCHCPMTYMKRKQCECRKRCNAGRECVLSEEDIRLK--LKQEEQAOATS 189
 Qy 132 ---RTGTPGVGLTEEQMMIRELMDAQMTFTDTSFHKFRPLPGVLSGCEPESL 188
 Db 190 VSPRVSSPPQVLPOLSEQIGMETKLVAAQOCNRRSFS----- 228
 Qy 189 QADREBAKWSQVRKDCSLKYSLOLRGEGSVWYKRPADSGGKEIFSLPHMADMT 248
 Db 229 ---DRLKVTWP-PIADP-----PQSRDARQORFA---HTELAI 260
 Qy 249 YMKGIISPAKVISYFRDLPIEDQISLKGAEFLCQLRPNYFNATGTCWECGRLSY 306
 Db 261 VSVQELVDPAKQVPGFLQLOREDOILAKASTIEMLLETARRNHET---ECIT 328
 Qy 307 CLEEDTA-GGFOQLLEPMLKFKHYMLKRLQLEHEEYVLMQALISLSPRPVGLQHRVVD 365
 Db 320 NREDPAKAGIQVEIFINPIEFSSAMNEDLNDAPRLLAIIFSSADRPNVDDQLOVEFL 379
 Qy 366 QEOPATLTKSYIECNRPQAPRFLFKIMAMTELRISNAHQRL--LRIDIHFPATP 423
 Db 380 QHTYVETALHAYVSNHPH--DRLMFRMLKLVSLRTLSVSHSEVAFALRLOD--KKLP 435
 Qy 424 LMOELFGI 431
 Db 436 LSEIMDV 443

RESULT 12
 138975
 nuclear orphan receptor LXR-alpha - human
 C:Species: Homo sapiens (man)
 C:Date: 16-Feb-1996 #sequence_revision 16-Feb-1996 #text_change 20-Sep-1999
 C:Accession: I38975


```

Db 74 GHELCNVCDDKASGFHYNYLSCEGKGFRRSVHVGAGRYACRSGGTCQDAFERRCQ 133
OY 96 ACRLRCLSGMCKKEMIMSDAEVEERBALIKRKSERGTOLP----- 138
Db 134 LCLRLRCKEAGMEQCVLSSEDIKRIKIOKQOQPPPTPAGSSARPAASPTSEAS 193
OY 139 -----GVQGLTEQRMRRLMDADMKTDTTFSHFKNRLPGVLSGCELPESLQA 190
Db 194 SGGSGEGEIO-LTAAQELMIQOLVAALQCKNRKRSF----- 229
OY 191 PSREAAKRSQYRKDCLSKVSLQJGDEGSSWYMKKPPADSGGEITSLPRHADMSTYM 250
Db 230 -DQPKTTPP-----LGAD-----PQSRDARQORFA--HFEELATIS 263
OY 251 FPGIISFAKVISYFRDLPEDQISLKGAFELCOLRNTYFNAETGWC-----GLSY 306
Db 264 VOEIVDFAKQVGFQJLGRDQIALKASTIEIMLEKARKYNET--ECITFLKDETY 320
OY 307 CLED-TAGGFQQLLEPLKFKHYMKLQJLHEEYVLMQALSLFSPDPGVLQHRVVDL 365
Db 321 SKDDEHRRAGLOVEFNPIFFESRAMRRLGLDAEYALLIAINIFRADRPVVOEPSREAL 380
OY 366 QEOFAITLKSXYECNRPOPARHFLKIMAMLTLSINAHQRL--LRIDIHPPATP 423
Db 381 QOPYEALLSTRTIKRPQDQLR--FPRIMKLVSLKTLSSVSHSQVAPALRLD--KKLP 436
OY 424 LMQELFGI 431
Db 437 LLESTMDV 444

```

RESULT 15

```

A:Accession: A56918
Species: Rattus norvegicus (Norway rat)
C:Date: 11-Aug-1995 #sequence_revision 11-Aug-1995 #text_change 24-Sep-1999
C:Accession: A56918
R:Forman, B.M.; Goode, E.; Chen, J.; Oro, A.E.; Bradley, D.J.; Perlmann, T.; Noonan, D.J.
Cell 81, 687-695, 1995
A:Title: Identification of a nuclear receptor that is activated by farnesol metabolites
A:Reference number: A56918; MVID:95292336
A:Accession: A56918
A:Status: preliminary; nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-469 <FOR>
A:Cross-references: GB:U18374; NID:9868031; PIDN:AC52205.1; PID:9868032
C:Superfamily: unassigned erda-related proteins; erda transforming protein homology
C:Keywords: DNA binding; nucleus; zinc finger
F:122-390/Domain: erda transforming protein homology <ERBA>

```

Query Match 19.9%; Score 456; DB 2; Length 469;

Best Local Similarity 28.6%; Pred. No. 5, 1e-28;

Matches 130; Conservative 78; Mismatches 146; Indels 100; Gaps 16;

```

OY 5 PKESWNADEFVCEDE---SVGKPSVNADE--EVGAPICRYCGDKATGYHNNWTC 58
Db 88 PDES-----VYQGETEVSMPYTKKPRMAASAGRIGDELCYVCGDRASGYHNNALTC 141
OY 59 EGCKGFERRAMRNARLCPFRKACETRTKTRQOCACRLKCLSGMCKKEMIMSDAV 118
Db 142 EGCKGFERSTKKNAYKCK-NGNCVMDMYRRKQDCRLKRCHEMGMLAECLLTIEIC 200
OY 119 EERRALIKRKSERTGTOLGVQ-----LTREQRMRLMDQMK 160
Db 201 KSKR-LRKVVKQHADQTVNDESGDLNQVSTYKLCREKTELTVDOITLDYIMDSYK 259
OY 161 TDDTFSHKRNRLPGVSSGCELPESLQAPSREAAKMSQVRKDLCLKVSJQLRGEDG 220
Db 260 Q-----RMPQET-----NKILKE----- 273
OY 221 SVWNYKPPADSGKEIFSLPRHADMSTYMKGIISPAKYISYFRDLPEDQISLKGAA 280

```

```

Db 274 -----EFSAEENFLIL--TEMATSHVQILVEFTKRLPGQTIDHEDQIALKSGA 321
OY 281 FELCOLRNTYFNAETGWEGRSLSYCLEDTAGCFQQLLEPLKFKHYMKLQJLHEEY 340
Db 322 VEMAFLRSEITPKKLPAGHADLLEERTKS--GISDEYITPMESFYKSVGELKMQEY 379
OY 341 VLMQALSLFSPDPGVLQHRVVDLOEQFAITLKSXYECNRPO-PARHFLKIMAMLT 399
Db 380 ALLTAIVILSPDQYIKDREAVEKQLPELDVLOKCKLKITOPENPH--FACLLGRTE 436
OY 400 LRSINAOHQRLL--RIODIHPPATPMLQELFGI 431
Db 437 LRTNHHHAEMLMMSRVND-HKF-TPLICEIMDV 468

```

Search completed: February 18, 2001, 14:33:07
Job time: 5356 sec

